Expectancies in Double-Blind Randomised Placebo-Controlled Trials and Placebo-Induced Side Effects

A thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in the Faculty of Science.

Ben Colagiuri
School of Psychology
University of Sydney

The research reported in these studies was approved by the following human research ethics committees, as appropriate:
University of Sydney
Central Sydney Area Health Service
South Eastern Sydney Area Health Service
Wentworth Area Health Service
University of Rochester (USA)
U.S. Department of Health and Human Services (USA)
Abstract

The majority of research on the placebo effect has focused on beneficial effects in patients or participants told to expect an active treatment, but who are actually given a placebo. Two important and relatively understudied aspects of the placebo effect are the extent to which expectancies influence outcomes in double-blind randomised placebo-controlled trials (RCTs) and whether the placebo effect contributes to treatment side effects. The current project investigated these two issues in both clinical and experimental settings. The first study involved reanalysing a double-blind RCT of naltrexone and acamprosate for alcohol dependence based on whether participants believed they had been allocated to receive active treatment or placebo (perceived treatment). The second study extended on this by developing an experimental model for these effects using dummy (placebo only) double-blind RCTs for cognitive performance. This allowed for the manipulation of observable changes in the form of false feedback. The third study investigated whether warning participants about side effects increases their occurrence, frequency, and/or severity in three dummy trials for sleep difficulty in healthy volunteers. The final study complemented this by examining whether first time chemotherapy patients’ expectancies for nausea were associated with their post-chemotherapy nausea. The studies on perceived treatment in double-blind RCTs indicated that participants’ beliefs about their treatment allocation can influence their actual treatment outcomes via the placebo effect and that these beliefs are affected by the feedback they receive about their performance. The studies on placebo-induced side effects indicated that the placebo effect may contribute to treatment side effects but that this effect is generally likely to be small. These findings confirm that the placebo effect can influence treatment outcomes and emphasise the importance of considering patient
expectancies when delivering medical treatment. They also highlight some general limitations associated with research on the placebo effect, which include, whether conveying uncertainty undermines the placebo effect and whether measuring or manipulating expectancies is the best way to evaluate the placebo effect.
Statement of Contribution

The research reported here is original and has not been presented for the award of any other degree. This thesis contains two experimental studies (Studies 2 and 3) and two analyses of pre-established studies (Studies 1 and 4). I was responsible for developing and implementing the experimental studies and did so in consultation with my supervisors. This involved 1) obtaining ethics approval, 2) developing the necessary information sheets, consent forms, and study specific questionnaires, 3) recruiting and testing all participants, 4) entering and analysing the data.

For the analyses of the two pre-established studies, data were kindly provided to me by Professor Paul Haber (Study 1) and Associate Professor Joseph Roscoe (Study 4). In both cases I was given complete discretion to choose how I analysed the data. As such, I was solely responsible for planning and conducting the analyses for these two studies. I have recently published the results of these analyses in peer-reviewed journals as first author. The co-authors for these publications contributed by their role in designing and implementing the original studies and/or commenting on drafts of the pre-publication manuscript. None of these co-authors have seen or commented on any version of this thesis, with the exception of Professor Robert Boakes, who is my Associate Supervisor.
Acknowledgements

Apart from a few moments of self doubt (and some more general moments of existential wonderment), I have genuinely enjoyed completing both my undergraduate degree in Psychology and now my PhD here at the University of Sydney. This is in no small part due to the support I have received during this time.

First and foremost I would like to thank my supervisors, Professors Robert Boakes and Phyllis Butow. I was fortunate enough to have Bob as my honours supervisor in 2005 and he is the main reason that I decided to stay in the School of Psychology to complete my PhD (winning out over Public Health). His love of research is inspiring, as is his academic integrity. He has been an outstanding supervisor in every way possible. Phyllis has also been a great supervisor. Despite her extremely busy schedule, she has always made time for me to discuss my research with her and even managed to read, comment on, and return a draft of one of the following chapters between me sending it to her, heading out to dinner, and then arriving home two or three hours later! Both Bob and Phyllis have been incredibly supportive of me and were very comfortable providing me with a high level of autonomy over my research, which I greatly appreciate.

I would also like to thank Professor Paul Haber, Dr Kirsten Morley and their colleagues from Drug Health Services, Royal Prince Alfred Hospital who were kind enough to allow me to reanalyse the data from their randomised controlled trial of naltrexone and acamprosate for alcohol dependence. Similarly, I would like to thank Associate Professor Joseph Roscoe and his colleagues from the University of
Rochester, USA for giving me free reign to analyse data on first time chemotherapy patients’ expectancies and subsequent nausea in a trial comparing antiemetic regiments for the treatment of delayed nausea. This was extremely generous of them.

My PhD was supported by an Australian Postgraduate Award provided by the Australian Government. I also received the Campbell Perry International Scholarship in 2007 which allowed me to visit and learn from researchers investigating the placebo effect in countries outside of Australia. I am very grateful to Mrs Jean Swirles for providing me with this opportunity as commemoration of her son, Emeritus Professor Campbell Perry, and to Mr Peter Doyle for his support of this scheme.

Finally, I would like to thank my family and friends for their support. My family has always been supportive of me and this has been no different for my PhD. We have also had two adorable additions to the family during my PhD, with my niece, Clare, being born in 2006 and my nephew, Josh, being born in 2008. Quite coincidently, Josh and Claire also happen to be the names of my office buddies and partners in crime for much of the past 3 or so years - a nice symmetry. Having Jesse, my International Collaborator, and Tjeerd in Australia for the last 18months has also been a great addition to my friends. And, as always, the Secret Seven et al. have been a great source of support (and distraction).
Table of Contents

Abstract ........................................................................................................................................... ii
Statement of Contribution ........................................................................................................ iv
Acknowledgements ................................................................................................................... v
Table of Contents .................................................................................................................. vii
List of Figures ........................................................................................................................ ix
List of Tables ........................................................................................................................... xi
List of Appendices .................................................................................................................. xii

Chapter 1 – Nomenclature, Evidence, and Theories of the Placebo Effect ................. 2
  1.1 Nomenclature .................................................................................................................... 2
  1.1.1 The placebo effect ..................................................................................................... 2
  1.1.2 Placebo-induced side effects .................................................................................. 5
  1.1.3 Expectancy .............................................................................................................. 5
  1.2 Evidence for the Placebo Effect ................................................................................... 6
  1.2.1 Anecdotal evidence for the placebo effect .............................................................. 6
  1.2.2 The importance of controlling for natural history ................................................... 7
  1.2.3 Scepticism about the placebo effect ....................................................................... 8
  1.2.4 Convincing evidence for the placebo effect .......................................................... 10
  1.3 Theories of the Placebo Effect ....................................................................................... 12
  1.3.1 Classical conditioning ............................................................................................ 12
  1.3.2 Expectancy theory .................................................................................................. 14
  1.3.3 Classical conditioning versus expectancy theory .................................................. 14
  1.3.4 Classical conditioning and expectancy theory: An integrated approach .......... 17
  1.4 Overview of the Current Project .................................................................................. 20

Chapter 2: Review of Expectancies in Double-Blind RCTs ........................................ 23
  2.1 Double-Blind Randomised Placebo-Controlled Trials .............................................. 23
  2.2 How Expectancies Can Limit the Validity of Double-Blind RCTs .............................. 25
    2.2.1 Failed blinding ....................................................................................................... 27
    2.2.2 Expectancies masking the active treatment’s effect ............................................. 33
    2.2.3 Differences between standard clinical practice and double-blind RCTs .......... 35
  2.3 Alternatives to double-blind RCTs .............................................................................. 40
    2.3.1 The Balanced Placebo Design ............................................................................. 41
    2.3.2 The ‘Open versus Hidden’ Design ...................................................................... 43
    2.3.3 Active rather than benign placebos ..................................................................... 44
  2.4 Conclusions .................................................................................................................... 46

Chapter 3: Perceived Treatment in a Double-Blind RCT for Alcohol Dependence (Study 1) .......................................................................................................................... 48
  3.1 Introduction ..................................................................................................................... 48
  3.2 Methods ........................................................................................................................ 49
  3.3 Results ............................................................................................................................ 53
  3.4 Discussion ...................................................................................................................... 60

Chapter 4 – Perceived Treatment, False Feedback, and Cognitive Performance: Two Dummy Double-Blind RCTs (Study 2) ................................................................. 63
  4.1 Introduction ..................................................................................................................... 63
  4.2 Experiment 1 ................................................................................................................ 65
  4.3 Methods ........................................................................................................................ 66
  4.4 Results and discussion ................................................................................................. 72
List of Figures

Chapter 1
Figure 1.1. Model of the mechanism of the placebo effect based on those proposed by Benedetti et al. (2003, p4321) and Stewart-Williams and Podd (2004, p336). 18

Chapter 2
Figure 2.1. The balanced placebo design. 41

Chapter 3
Figure 3.1. Covariate adjusted mean (+SE) total alcohol consumption by actual and perceived treatment. 55
Figure 3.2. Covariate adjusted mean (+SE) alcohol dependence score (ADS) across actual and perceived treatment. 55
Figure 3.3. Covariate adjusted mean (+SE) alcohol cravings (PACS). 55
Figure 3.4. Covariate adjusted mean (+SE) number of days abstinent across actual and perceived treatment. 56
Figure 3.5. Covariate adjusted mean (+SE) number of days until relapse across actual and perceived treatment allocation. 57
Figure 3.6. Covariate adjusted mean (+SE) number of adverse side effects across actual and perceived treatment allocation. 59

Chapter 4
Figure 4.1. Design for Experiment 1. 67
Figure 4.2. Experiment 1. Feedback on forced choice (FC) perceived treatment. 72
Figure 4.3. Experiment 1. Mean (±SE) accuracy (A) and reaction time (B) for those who received positive feedback and those who received no change feedback. 73
Figure 4.4. Experiment 1. Mean (±SE) accuracy (A) and reaction time (B) based on participants forced choice perceived treatment. 74
Figure 4.5. Design for Experiment 2. 78
Figure 4.6. Experiment 2. Feedback on forced choice (FC) perceived treatment. 81
Figure 4.7. Experiment 2. Mean (±SE) accuracy (A) and reaction time (B) for those who received positive feedback and those who received no change feedback. 82
Figure 4.8. Experiment 2. Mean (±SE) accuracy (A) and reaction time (B) based on participants forced choice perceived treatment. 83
Figure 4.9 Experiment 2. Mean (+SE) accuracy (A) and reaction time (B) on the post-feedback test by feedback for those who were in the double-blind group and those in the no treatment group. 85

Chapter 6
Figure 6.1. Experiment 1. Mean (+SE) global sleep difficulty across treatment. 120
Figure 6.2. Experiment 1. Mean (+SE) sleep latency (A) and total sleep time across groups. 121
Figure 6.3. Experiment 1. Occurrence (%) of target symptoms across treatment groups. 123
Figure 6.5. Experiment 1. Mean (+SE) severity rating for target symptoms across treatment condition. 125
Figure 6.6: Experiment 2. Mean (+SE) global sleep difficulty before and after treatment for each group. 130
Figure 6.7. Experiment 2. Mean (+SE) sleep latency (A) total sleep time (B) and by treatment group..........................................................131
Figure 6.8: Experiment 2. Mean (+SE) occurrence (total number of symptoms, A), frequency (B), and severity (C) of target symptoms versus non-target symptoms reported by the suggestion group..........................................................133
Figure 6.9. Experiment 2. Occurrence (A) and mean (+SE) frequency (B) and severity (C) of each symptom in the suggestion group when suggested (target), in the suggestion group when not suggested (non-target), and in the control group........134
Figure 6.10. Experiment 2. Correct recall of side effects contained in the warning given to participants in the suggestion group. ..........................................................135
Figure 6.11. Experiment 3. Mean (+SE) global sleep difficulty (A) and total sleep time before and after treatment for the suggestion and control group ..................140
Figure 6.12. Experiment 3. Proportion and direction of changes in appetite for participants in the suggestion and control groups.................................142
Figure 6.13. Experiment 3. Mean (+SE) severity of changes in appetite for participants in the suggestion and control groups.................................142

Chapter 7

Figure 7.1. Covariate adjusted mean (+SE) average (A) and peak (B) nausea by level of expectancy for all participants..........................................................159
List of Tables

Chapter 3
Table 3.1. Baseline characteristics across actual treatment and perceived treatment. Means (SE) are given for continuous variables, percentage male is given for sex, and * denotes \( p < .05 \) ........................................................................................................... 53
Table 3.2. Distribution of participants according to their actual treatment allocation and their perceived treatment allocation. ........................................................................ 54
Table 3.3. Mean (SE) number of days compliant with treatment instructions (compliance), number of counselling sessions attended (counselling), and posttreatment motivation to change (SOCRATES) by actual and perceived treatment. No differences were significant. .................................................................................. 59

Chapter 5
Table 5.1. Summary of studies assessing the relationship between first time chemotherapy patients’ expectancies and post-chemotherapy nausea. In all studies expectancies were assessed prior to the first infusion. In each case the patients, statistical analysis, measure of nausea, significance of expectancy, and limitations are described. Where more than one statistical test was used, multivariate analysis was favoured over univariate analysis. Other predictors included (hierarchical) or entered (stepwise) into the regression models that significantly predicted nausea are italicised. If provided by the authors, the unique proportion of variability (\( \Delta R^2 \)) expectancy accounted for in post-chemotherapy nausea is given. ..................................................... 101

Chapter 6
Table 6.1. Design of Experiment 1. Participants were allocated to receive placebo treatment or no treatment (control). Half of those allocated to receive placebo treatment were warned about three possible side effects (suggestion) while the other half were not (no suggestion)………………………………………………………115

Chapter 7
Table 7.1. Final step in the hierarchical regression of predictors of post-chemotherapy nausea for all participants. †Dummy coded with breast cancer as the reference group. ††Dummy coded with Study Arm 1 as the reference group *Significant at \( p < .05 \), **Significant at \( p < .01 \), ***Significant at \( p < .001 \). ........................................................................................................ 156
Table 7.2. Simultaneous regression of predictors of expectancy for nausea. †Dummy coded with breast cancer as the reference group. ***Significant at \( p < .001 \). .......... 161
List of Appendices

Appendix 1 – List of published and submitted peer-reviewed journal articles based on work undertaken for the degree of Doctor of Philosophy.......................................................... 193

Appendix 2 – List of conference presentations based on work undertaken for the degree of Doctor of Philosophy ........................................................................................................ 194

Appendix 3a – Information Sheet for Study 2, Experiment 1. ................................................................. 195
Appendix 3b – Beliefs Questionnaire for Study 2, Experiment 1. ............................................................. 197
Appendix 3c – Information Sheet for Study 2, Experiment 2. (Double-blind) ..................................... 199
Appendix 3d – Information Sheet for Study 2, Experiment 2. (Control Group) .................................... 201
Appendix 3e – Beliefs Questionnaire for Study 2, Experiment 2. (Double-blind) .............................. 203
Appendix 3f – Beliefs Questionnaire for Study 2, Experiment 2. (Control Group) .............................. 205

Appendix 4a – Information Sheet for Study 3, Experiment 1 (Suggestion) ........................................ 206
Appendix 4b – Posttreatment Questionnaire for Study 3, Experiment 1 ........................................... 208
Appendix 4c – Adverse Symptom Checklist for Study 3, Experiment 1 ............................................. 209
Appendix 4d – Information Sheet for Study 3, Experiment 2 (Suggestion A) ..................................... 211
Appendix 4e – Information Sheet for Study 3, Experiment 2 (Suggestion B) ..................................... 213
Appendix 4f – Posttreatment Questionnaire for Study 3, Experiment 2 ........................................... 215
Appendix 4g – Adverse Symptom Checklist for Study 3, Experiment 2 ........................................... 216
Appendix 4h – Information Sheet for Study 3, Experiment 3 (Increase) ............................................. 218
Appendix 4i – Information Sheet for Study 3, Experiment 3 (Decrease) ............................................ 220
Appendix 4j – Posttreatment Questionnaire for Study 3, Experiment 3 ........................................... 222

Appendix 1 - List of published and submitted peer-reviewed journal articles based on work undertaken for the degree of Doctor of Philosophy.......................................................... 193
Appendix 2 – List of conference presentations based on work undertaken for the degree of Doctor of Philosophy ........................................................................................................ 194
Appendix 3a – Information Sheet for Study 2, Experiment 1. ................................................................. 195
Appendix 3b – Beliefs Questionnaire for Study 2, Experiment 1. ............................................................. 197
Appendix 3c – Information Sheet for Study 2, Experiment 2. (Double-blind) ..................................... 199
Appendix 3d – Information Sheet for Study 2, Experiment 2. (Control Group) .................................... 201
Appendix 3e – Beliefs Questionnaire for Study 2, Experiment 2. (Double-blind) .............................. 203
Appendix 3f – Beliefs Questionnaire for Study 2, Experiment 2. (Control Group) .............................. 205

Appendix 4a – Information Sheet for Study 3, Experiment 1 (Suggestion) ........................................ 206
Appendix 4b – Posttreatment Questionnaire for Study 3, Experiment 1 ........................................... 208
Appendix 4c – Adverse Symptom Checklist for Study 3, Experiment 1 ............................................. 209
Appendix 4d – Information Sheet for Study 3, Experiment 2 (Suggestion A) ..................................... 211
Appendix 4e – Information Sheet for Study 3, Experiment 2 (Suggestion B) ..................................... 213
Appendix 4f – Posttreatment Questionnaire for Study 3, Experiment 2 ........................................... 215
Appendix 4g – Adverse Symptom Checklist for Study 3, Experiment 2 ........................................... 216
Appendix 4h – Information Sheet for Study 3, Experiment 3 (Increase) ............................................. 218
Appendix 4i – Information Sheet for Study 3, Experiment 3 (Decrease) ............................................ 220
Appendix 4j – Posttreatment Questionnaire for Study 3, Experiment 3 ........................................... 222

Appendix 4k – Adverse Symptom Checklist for Study 3, Experiment 3 ........................................... 223
“One of the most successful physicians I have ever known has assured me that he used more bread pills, drops of colored water, and powders of hickory ashes than of all other medicines put together. It was certainly a pious fraud.”

Thomas Jefferson, 1807
Chapter 1 – Nomenclature, Evidence, and Theories of the Placebo Effect

Medical treatment involves much more than the direct physiological action of the substance or procedure being administered. Treatment is typically administered or suggested by a health professional who communicates – intentionally or otherwise - the aim of the treatment, how likely it is to be effective, and whether or not it might produce side effects. The treatment occurs within a context, for example, a hospital ward, a general practice surgery, or at home. The way one treatment is delivered often differs from another, both in terms of invasiveness and complexity of the treatment regimen. In addition to this, patients often have preconceptions about the likely effects of a treatment based on previous experience, the experience of others, or what they have read or heard in the media. All of these factors are likely to affect what a patient expects from his/her treatment and these expectancies can influence the patient’s actual health outcomes, either positively or negatively, via the placebo effect. Understanding the placebo effect, then, may enable health professionals to maximise beneficial treatment outcomes whilst minimising adverse outcomes. This chapter reviews nomenclature, evidence and theories relating to the placebo effect before providing an overview of the current project.

1.1 Nomenclature

1.1.1 The placebo effect

The current literature contains varied definitions of the placebo effect. In most cases the differences do not reflect any serious disagreement about what constitutes a placebo effect, but are, instead, testament to the complexity of the placebo effect. The
only substantial difference concerns whether the placebo effect should be confined to positive responses or whether it should also include negative responses.

Researchers who define the placebo effect as exclusively positive often focus on clinical settings and refer to the therapeutic benefits experienced by the patient (e.g. Brody & Brody, 2000; Shapiro & Shapiro, 1997a, 1997b). This generally leads to separation of positive and negative responses, labelled respectively as the placebo effect and the nocebo effect (e.g. Evans & Rogers, 2003; Hahn, 1997a, 1997b; Harrington, 1999). According to these definitions, the placebo effect is when expectancies for positive outcomes cause positive responses and the nocebo effect is when expectancies for negative outcomes cause negative responses. While discriminating between positive and negative placebo effects can be quite useful for labelling such effects, separating them into placebo effects and nocebo effects leads to a number of logical inconsistencies.

Firstly, as with many medications and procedures, placebos can produce both beneficial and adverse effects simultaneously (e.g. Downing, Rickels, Rickels, & Downing, 1979). Therefore, defining these effects separately as placebo effects and nocebo effects would lead to the potentially confusing conclusion that a substance or procedure could be both a placebo and a nocebo at the same time (Siegel, 2002; Stewart-Williams & Podd, 2004). Secondly, it may not always be clear whether a response is positive or negative (Stewart-Williams & Podd, 2004). For example, placebo alcohol could cause a feeling of intoxication, which one person might experience as pleasant, while another person experiences it as unpleasant. This would mean that the placebo alcohol produced a placebo effect in the first person but a
nocebo effect in the second person. This is potentially confusing because the only
difference between the placebo effect and nocebo effect here would be the
individuals’ hedonic experiences, not the actual response, which was the same in both
cases, i.e. intoxication. Thirdly, some responses may be neutral, especially if they are
undetectable. For instance, minor increases or decreases in heart rate caused by
placebo coffee are unlikely to be experienced as either positive or negative. If placebo
effects are only positive and nocebo effects are only negative, then neutral responses
must, rather illogically, be neither of these.

Stewart-Williams and Podd (2004) have provided a more general definition of
a placebo effect which appears to overcome these problems. They state that “a
placebo effect is a genuine psychological or physiological effect, in a human or
another animal, which is attributable to receiving a substance or undergoing a
procedure but is not due to the inherent powers of that substance or procedure” (p.
326). In this way, the placebo effect is not confined to positive responses but includes
positive, negative, and even neutral responses. Further, the genuineness of the placebo
effect indicates that it is a real effect which cannot be accounted for by biases, such
as, demand characteristics. The fact that it is attributable to receiving a treatment or
procedure means that it is an effect beyond natural history, spontaneous remission,
and regression to the mean. By extension, a placebo is simply a substance or
procedure that has no inherent power to affect the psychological or physiological
process of interest.
1.1.2 Placebo-induced side effects

Side effects are generally defined as responses to a treatment other than those for which the treatment is being administered (Barsky, Saintfort, Rogers, & Borus, 2002). These can be beneficial, adverse, or neutral. An example of an adverse side effect is the gastrointestinal discomfort that sometimes follows the administration of aspirin for pain relief. If a placebo pill was administered instead of aspirin and the recipient also experienced pain relief and gastrointestinal discomfort, then, for the reasons stated above, both the pain relief and the gastrointestinal discomfort should be considered placebo effects. There is, however, some benefit in labelling these two responses differently. The pain relief caused by the placebo pill can simply be labelled a placebo effect, while the gastrointestinal discomfort it produced can be labelled a placebo-induced side effect. Although there may be no difference in the aetiology of these two effects, the placebo pill was administered for pain relief, whereas the gastrointestinal discomfort was a side effect of its administration. Importantly, this is identical to the distinction made between non-placebo treatment effects and their side effects. Pain relief caused by aspirin is aspirin’s treatment effect because it was administered for that purpose and the gastrointestinal discomfort is aspirin’s side effect. In this way, labelling responses as placebo effects and placebo-induced side effects appropriately parallels discussion of non-placebo effects.

1.1.3 Expectancy

Expectancy is an intuitive although somewhat vague concept. It refers to a subjective sense of the probability of a future event. A number of different terms are used for expectancy in relation to placebo effects, which are often used interchangeably. These include: belief, conscious expectancy, and response
expectancy. The choice of the term used in this thesis will attempt to highlight some relevant characteristic of the expectancy being discussed. For example, when comparing expectancy and classical conditioning accounts of the placebo effect I will refer to conscious expectancies to emphasise that awareness reflects an important difference between these two theories. When discussing expectancies in double-blind RCTs I will refer to beliefs about treatment allocation so as to contrast this from expectancies regarding the efficacy of a treatment, even though they are both forms of expectancy that may contribute to the placebo effect.

1.2 Evidence for the Placebo Effect

1.2.1 Anecdotal evidence for the placebo effect

Some interesting anecdotal evidence for the placebo effect has emerged over the past few centuries. Shapiro and Shapiro (1997a; 1997b), for example, attribute most of the efficacy of pre-scientific medicine to the placebo effect. In the extreme, Benson (1997) cites voodoo death as an example of a negative placebo effect. Perhaps the most fascinating anecdotal account of the placebo effect is that of Mr Wright (Klopfer, 1957). Mr Wright was suffering from cancer of the lymph nodes and had not responded to any of the available anti-cancer treatments. As a result, both he and his physicians expected that he would not recover from the cancer. Mr Wright then heard about a new anti-cancer drug called krebiozen and became convinced that this would cure his cancer. After unrelenting requests, his physician injected him with krebiozen. Almost immediately, Mr Wright’s tumours rapidly reduced in size, so much so that he felt cured. That is, until he read preliminary findings that indicated that krebiozen was ineffective, whereupon he relapsed as quickly as he had recovered. His physician attempted to reverse this by telling him that the preliminary results
regarding krebiozen were unconvincing and promised to inject him with a more potent dose of krebiozen. The injection actually contained water, but nonetheless Mr Wright again showed remarkable improvement. Sometime after this, Mr Wright read a final report declaring that krebiozen was completely ineffective. He relapsed and died two days later.

1.2.2 The importance of controlling for natural history

Perhaps not surprisingly, these types of stories have led some investigators to claim that the placebo effect seems capable of curing any condition (e.g. Buckman & Sabbagh, 1993). There is, however, one crucial aspect of health and illness that these stories do not control for, that is, natural history. Natural history refers to the course of a condition when untreated and involves the duration of the condition, spontaneous remission or recovery, and random fluctuation in symptoms. Although medical treatments can effectively reduce the severity and duration of an illness, in many cases the illness will subside without treatment. For example, most people experiencing the common cold recover after six days simply as a result of normal immune-functioning (Diehl, 1953). In other cases, illnesses can spontaneously remit either temporarily or permanently without apparent explanation, as is common in cancer patients (see Huebscher, 1992 for a review). In addition to this, symptoms of both health and illness tend to fluctuate randomly, as evidenced by variations in pain intensity experienced by people with chronic pain (e.g. Foss, Apkarian, & Chialvo, 2006). These naturally occurring effects can be mistaken for placebo effects when placebo administration precedes a change in symptoms that they produce. For example, administering a placebo when pain is at its greatest is very likely to correspond with a reduction in pain due to the natural progression of pain, not the placebo effect.
(Benedetti, 2009; Colloca & Benedetti, 2005). In order to detect a real placebo effect, then, it is necessary to compare placebo administration with a natural history group who receive no treatment, something which the anecdotal evidence above clearly lacks.

1.2.3 Scepticism about the placebo effect

In stark contrast to anecdotal stories of extreme placebo effects and subsequent claims that the placebo effect is all powerful, some researchers have questioned whether the placebo effect exists at all. Most notably, Hróbjartsson and Gøtzsche (2001; 2004) reviewed double-blind RCTs that included both a placebo group and a no treatment group and concluded that there was no evidence that the placebo effect was clinically significant. In their first review, they identified 114 double-blind RCTs and categorised them according to the type of outcome measure used; either dichotomous or continuous, and objective or subjective. When analysing these pooled data they only found a significant placebo effect for studies with continuous, subjective outcome measures and for trials on pain and these effects were only small. In their second review, they conducted an identical analysis but added 52 more recent double-blind RCTs that included a no treatment group. Again, they found a significant placebo effect only for studies with continuous, subjective outcome measures but this time they found these effects for pain and phobias. However, they considered the significant result for phobia unreliable due to small sample size (n=57) and argued that the effect size for pain, which corresponded to a 6mm decrease on a 100mm visual analogue scale, was unlikely to be clinically significant. Based on these reviews, they concluded that the placebo effect is generally grossly overestimated and has little clinical importance.
There are however, two important limitations to Hróbjartsson and Gøtzsche’s (2001; 2004) reviews which suggest that they underestimated the magnitude of the placebo effect. Firstly, they analysed the difference between placebo treatment and no treatment in double-blind RCTs. As discussed in detail in the next chapter, double-blind RCTs often detract from the placebo effect because participants in these trials are aware that they may be receiving a placebo. That is, the placebo groups in their reviews consisted of patients given placebo treatment who may or may not have expected that they were receiving an active treatment, rather than patients given placebo treatment and who are told that they are receiving active treatment. As such, smaller placebo effects would be expected in the placebo groups they analysed. Vase, Riley, Price, and Vase (2002) demonstrated this by showing that the placebo effect was much larger in studies examining placebo analgesia compared with the double-blind RCTs on pain included in Hróbjartsson and Gøtzsche’s (2001) first review.

Secondly, Hróbjartsson and Gøtzsche’s (2001; 2004) reviews combined data from over forty different maladies. The pain studies, for instance, included trials on headaches, rheumatoid arthritis, low back pain, and fibromyalgia. This ignores the possibility that some conditions may be more amenable to the placebo effect than others (Evans, 2003) or even that some types of placebos might be more effective for one condition than another. Wampold, Minami, Tierney, Baskin, and Bhati (2005) reanalysed the trials included in Hróbjartsson and Gøtzsche’s (2001) review, but also categorised each condition as amenable to the placebo effect or not. This was done by asking five doctoral psychology students who were unaware of Wampold et al.’s (2006) aims to rate the likelihood that each condition could be affected by the placebo effect. The results indicated a large placebo effect for conditions classified as
amenable and no placebo effect for conditions classified as non-amenable. In light of these limitations, Hróbjartsson and Gøtzsche’s (2001; 2004) reviews appear inadequate to estimate the true magnitude of the placebo effect. In order to achieve this, well controlled studies that investigate the placebo effect on specific conditions are required.

1.2.4 Convincing evidence for the placebo effect

There are a number of well controlled experimental studies that provide convincing evidence for the placebo effect in both clinical and non-clinical settings. The vast majority of these are on pain and typically show that a placebo, such as a benign topical cream or injection, believed to be an analgesic, produces pain relief that cannot be explained by natural history (e.g. Amanzio & Benedetti, 1999; Benedetti, Arduino, & Amanzio, 1999; Benedetti, Pollo et al., 2003; Montgomery & Kirsch, 1997; Voudouris, Peck, & Coleman, 1985; Voudouris, Peck, & Coleman, 1989, 1990). Some of these studies are described in detail later in this chapter. An important aspect of placebo analgesia is that it can be reversed by a hidden injection of naloxone (Amanzio & Benedetti, 1999; Benedetti, Arduino et al., 1999; Levine & Gordon, 1984; Levine, Gordon, & Fields, 1978), suggesting that there is a physiological corollary to the pain relief and that the effect exists beyond demand characteristics. This point is strengthened by Zubeita et al.’s (2005) finding that participants exhibiting placebo analgesia showed greater activation of the endogenous opioid system, which inhibits pain and stress via μ-opioid receptors, than participants who did not show a placebo effect.
There are also a handful of well controlled studies on the placebo effect in clinical settings other than pain. Placebo administration appears to reduce the sleep difficulty associated with sleeping in a laboratory for the first night compared with when no placebo is administered (Suetsugi, Mizuki, Yamamoto, Uchida, & Watanabe, 2007). People with Parkinson’s disease exhibit better motor performance when they are aware that they are receiving stimulation of the subthalamic nucleus compared with when they are unaware (Benedetti, Maggi et al., 2003; Mercado et al., 2006; Pollo et al., 2002). Smokers attempting to quit have lower withdrawal symptoms when they are told they are receiving nicotine compared with when they are told they are receiving a placebo, regardless of whether or not they actually receive nicotine (Gottlieb, Killen, Marlatt, & Taylor, 1987). In addition, there is indirect evidence for the placebo effect in clinical settings, in that more invasive treatment regimens appear to elicit larger placebo effects. For example, placebo injections reduced pain due to migraine headaches more effectively than did placebo pills (de Craen, Tijssen, de Gans, & Kleijnen, 2000) and four placebo pills per day shortened recovery rates for stomach ulcers more than two placebo pills per day (de Craen et al., 1999).

In terms of non-clinical settings, decaffeinated coffee has been shown to improve reaction times in sleep deprived healthy volunteers when they are told that it contains caffeine compared with when they are aware that it is decaffeinated (Anderson & Horne, 2008). Healthy volunteers told that they are being given a relaxant or a stimulant, but are in fact given a placebo, show responses in the direction of the information they received, and this information can moderate the response to an active drug (Flaten et al., 1991). Healthy volunteers who expect that alcohol will
impair their cognitive performance demonstrate poorer information processing than those who expect that alcohol will not affect their cognitive performance, both when alcohol and placebo alcohol are administered (Fillmore, Carscadden, & Vogel-Sprott, 1998). Taken together, these studies provide firm evidence for the placebo effect as measured by self report as well as when measured objectively.

1.3 Theories of the Placebo Effect

Traditionally, two theories have been relied upon to account for the placebo effect: classical conditioning and expectancy theory. While these two accounts were often treated as mutually exclusive, most researchers now agree that both classical conditioning and conscious expectancy can contribute to the placebo effect (Brody & Brody, 2000; Evans, 2003; Pacheco-López, Engler, Niemi, & Schedlowski, 2006; Stewart-Williams, 2004; Stewart-Williams & Podd, 2004). This is a result of evidence that neither account, on its own, can explain the full range of placebo effects found in humans and other animals and the acknowledgement that classical conditioning can produce conscious expectancies.

1.3.1 Classical conditioning

Pavlov (1927) showed that by pairing a tone with the delivery of morphine, dogs became restless and began to salivate when they heard the tone, even without the morphine itself. He described this processes as follows. The morphine was an unconditioned stimulus (US) in that it unconditionally caused the dogs to become restless and salivate. The dogs’ restlessness and salivation was an unconditioned response (UR), because it was caused unconditionally by the morphine. The tone, which was initially neutral with respect to the dogs’ activity and salivation, functioned
as a conditioned stimulus (CS). Through repeated pairings with the morphine, it acquired the power to produce restlessness and salivation in the dogs even in morphine’s absence. The restlessness and salivation in response to the tone was a conditioned response (CR) because it was conditional on the pairing of the tone and the morphine. Pavlov (1927) proposed that through repeated CS-US pairings the CS acquired the power to produce an effect, the CR, which was previously reserved for the US. This process of learning is known as the stimulus substitution model of classical conditioning.

Hernstein (1962) noticed how similar the stimulus substitution model of classical conditioning was to those that produce placebo effects. He demonstrated that after conditioning a rat with injections of scopolamine hydrobromide, which suppresses learned behaviour, an injection of saline also began to suppress the rat’s learned behaviour. He concluded that placebo effects must be instances of classical conditioning. Wickramasekera (1980) expanded on this. He proposed that when a person undergoes treatment, the context in which it is delivered, including the route of administration, the surroundings in which it is delivered, and even the person administering it, acts as a complex CS. This complex CS, through repeated pairings with the US, i.e. the drug itself, acquires the power to produce the drug effect on its own accord. In this way, proponents of the classical conditioning account argued that the placebo effect is simply a conditioned response that develops through CS-US pairings.
1.3.2 Expectancy theory

Kirsch (1985; 1997; 1999), on the other hand, has proposed an account of the placebo effect based on response expectancies. Response expectancies are anticipations of nonvolitional responses, such as, mood states, fear and anxiety, sexual arousal, perceptions of pain, and asthmatic responses (Kirsch, 1999). These response expectancies can be produced in a number of ways, including both direct and vicarious experience and verbal information (Kirsch, 1997). According to Kirsch, the anticipation of a response is sufficient to produce the response in and of itself. That is, activation of a response expectancy directly elicits that response. If a person expects pain relief after medical treatment, they will indeed experience pain relief. Similarly, a person who expects to feel alert after drinking a cup of coffee will feel alert after drinking it. Further, he viewed response expectancies as self-confirming. Once a response expectancy has been established it is sufficient to continue producing the response (Kirsch, 1997). In this way, he argued that placebo effects are the direct result of response expectancies. Kirsch (1985) was also adamant that this process is not confined to subjective states, but includes both the physiological processes that underlie them and physiological processes not related to subjective states. Of course, response expectancies are not the only determinants of behaviour. Rather, behaviour is the result of the interaction between conscious expectancies and other factors, such as the drug effect itself (Kirsch, 1997).

1.3.3 Classical conditioning versus expectancy theory

A series of studies conducted from the mid-1980’s to the late-1990’s attempted to determine whether classical conditioning or expectancy theory could explain the placebo effect. Voudouris, Peck and Coleman (1985; 1989; 1990) showed
that the direction of the placebo effect followed the direction of the conditioning
participants received, even when it contradicted the verbal information provided.
They told two groups of participants that a placebo cream was an analgesic and gave
one group surreptitious conditioning supporting this, i.e. lower pain intensity with the
cream, while the other group received surreptitious conditioning contradicting this,
i.e. higher pain intensity with the cream. Participants who received conditioning
supporting the verbal suggestion that the cream was an analgesic showed placebo
analgesia during the test phase, while those who received conditioning contradicting
the verbal suggestion showed placebo hyperalgesia. As a result, Voudouris et al.
(1985; 1989; 1990) argued that classical conditioning causes placebo effects, not
conscious expectancy.

Montgomery and Kirsch (1997) countered this by arguing that in the group
that received contradictory conditioning, the conditioning trials simply created
conscious expectancies that were stronger than the verbal information the participants
received. Using a similar design to Voudouris et al. (1985; 1989; 1990), they
administered placebo cream under the guise of an analgesic to two groups of
participants. Both of these groups received conditioning trials consisting of reduced
pain stimulation when the cream was applied. For one group the conditioning was
surreptitious, as in the original study, however, the second group were informed that
the painful stimulation was being decreased on the conditioning trials when the cream
was applied. Despite the fact that both groups received identical conditioning, only
those who were unaware of the conditioning manipulation exhibited placebo
analgesia. That is, being informed that the conditioning was taking place eliminated
the placebo effect. As a result, they concluded that the responses observed must have
been the result of the conscious expectancies created by interpretation of the conditioning trials, not the conditioning itself.

However, as Montgomery and Kirsch (1997) admit, their findings are only inconsistent with the stimulus substitution model of classical conditioning. Recent advances in classical conditioning theory have extended it from the low-order biological reflex first described by Pavlov to a more sophisticated learning process through which the organism comes to understand its environment (Rescorla, 1988). Rescorla (1988) argues that the importance of a conditioning trial lies in the information it provides about how one event relates to another. As such, simple CS-US pairings may not be sufficient to produce learning. Instead, these pairings must convey some form of information about the structure of events. Thus, while the stimulus substitution model cannot account for the above findings, a contemporary learning theorist would argue that the awareness regarding the conditioning manipulation devalued the CS-US pairings and hence produced no learning (Evans, 2003; Stewart-Williams & Podd, 2004). This means that both classical conditioning and expectancy theory can explain Voudouris et al. (1985; 1989; 1990) and Montgomery and Kirsch’s (1997) findings.

Nevertheless, there is evidence to suggest that neither account, on its own, can explain all types of placebo effect. Expectancy theory posits that any conditioning leading to placebo effects must be consciously mediated (Kirsch, 1997; Montgomery & Kirsch, 1997) and, as a result, that all placebo effects must be consciously mediated. There is, however, evidence from the conditioning literature that conditioned responding can be dissociated from conscious expectancies (c.f. the
Perruchet effect: Perruchet, 1985; Perruchet, Cleeremans, & Destrebecqz, 2006; Weidemann, Tangen, Lovibond, & Mitchell, 2009), which suggests that a placebo effect could occur either without the participant’s awareness or even in contrast to the participant’s expectancies. Benedetti, Amanzio, Baldi, Cassadio and Maggi (1999) have provided evidence that supports this. They found that a placebo led to respiratory depression after conditioning with buprenorphine, a mild respiratory depressant, in the absence of verbal suggestion and even though participants did not notice this response during conditioning or placebo administration. On the other hand, there is evidence that placebo effects can occur without prior drug experience (Pihl & Altman, 1971). For example, Wolf (1950) reported that information that ipecac had anti-emetic properties reversed the nausea a young woman was experiencing, even though this woman was unlikely to have had prior experience with the substance. Classical conditioning has difficulty accounting for these types of placebo effects because there has been no opportunity for conditioning to occur. Expectancy theory can, however, explain these types of placebo effects because verbal information is a source of response expectancies and, as such, is sufficient to produce a placebo effect. Thus, while both classical conditioning and conscious expectancies theory are sufficient to produce a placebo effect, neither appear necessary.

1.3.4 Classical conditioning and expectancy theory: An integrated approach

Almost simultaneously, Benedetti, Pollo et al. (2003) and Stewart-Williams and Podd (2004) developed similar integrative models of the placebo effect. An interpretation of these two models is presented in Figure 1.1. It shows that information and classical conditioning produce conscious expectancies which lead to consciously mediated placebo effects. However, it also allows for nonconsciously
mediated placebo effects that are the direct result of classical conditioning. As such, consciously mediated placebo effects can be explained by what the person expects to happen, i.e. their conscious expectancy, which could result from information or prior experience. Nonconsciously mediated placebo effects, on the other hand, can occur in the absence of conscious expectancy or even contrary to it. These effects cannot be explained by what the person expects to happen, but rather by his/her previous experience. In this way, this model can account for evidence that verbal suggestion without prior experience leads to a placebo effect because verbal information is sufficient to produce a conscious expectancy and thereby a placebo effect. It can also account for placebo effects which participants appear unaware of, because it allows for classically conditioned placebo effects that are not consciously mediated.

![Model of the mechanism of the placebo effect based on those proposed by Benedetti, Pollo et al. (2003, p4321) and Stewart-Williams and Podd (2004, p336).](image)

*Figure 1.1.* Model of the mechanism of the placebo effect based on those proposed by Benedetti, Pollo et al. (2003, p4321) and Stewart-Williams and Podd (2004, p336).

Support for this model comes from two studies conducted by Benedetti and his colleagues. In the first, Amanzio and Benedetti (1999) partitioned the contribution of information-based conscious expectancy and conditioning without conscious expectancy to placebo analgesia. To do this, they gave one group of participants a placebo injection (saline) but told them that it was a powerful analgesic, representing information-based conscious expectancy. A second group received prior conditioning
with morphine but were told that the placebo injection given in the test phase was an antibiotic that would have no effect on pain sensitivity. This group reflects conditioning without conscious expectancy, because they were told that the placebo injection was different from the one which they had been conditioned with. A third group received conditioning with morphine and were told that the placebo injection was the same powerful analgesic. This group represents a potential combination of conscious expectancy and nonconscious conditioning, because they were told the placebo injection was an analgesic and had received conditioning supporting this. In the test phase, all groups displayed significant placebo analgesia compared with a natural history group, suggesting that both information-based conscious expectancy and conditioning without conscious expectancy can produce placebo effects. Importantly, placebo analgesia was larger in the combined group than in the other two groups, indicating that neither information-based conscious expectancy nor conditioning without conscious expectancy could fully account for this effect. Instead, it would seem that these two mechanisms combined to produce a placebo effect greater in magnitude than either could on its own.

In the second study, Benedetti, Pollo et al. (2003) showed that information affects consciously mediated placebo effects but not nonconsciously mediated placebo effects. The consciously mediated placebo effects they examined were for pain in healthy participants and motor performance in Parkinson’s disease. In these conditions, the placebo effect followed the direction of the verbal suggestions participants received, even when it contradicted the conditioning they had received. For example, healthy participants who were told that a placebo injection was an analgesic showed placebo analgesia, while those who were told that a placebo
injection was a hyperalgesic showed placebo hyperalgesia, even though both groups had received prior conditioning with analgesic injections. The nonconsciously mediated placebo effects were for hormonal secretion. In these conditions information had no effect on hormone levels but conditioning did. For example, after conditioning with sumatriptan, growth hormone levels increased in response to placebo administration, mimicking the drug’s effect, regardless of whether participants were told that the placebo injection would increase or decrease their growth hormone levels. These two studies, therefore, support an integrative model of the placebo effect that incorporates the interaction between information and classical conditioning in producing conscious expectancies, as well as allowing for nonconscious placebo effects that result directly from classical conditioning. Thus, in exploring the placebo effect, it is important to consider both the information the participant has received and any prior experience he or she has had with the treatment itself and/or the treatment setting.

1.4 Overview of the Current Project

Traditionally, the primary focus of research on the placebo effect has been on beneficial effects in patients/participants told to expect an active treatment but who are actually given a placebo. More recently, there has been interest in whether participant expectancies affect double-blind RCTs and whether the placebo effect contributes to treatment-related side effects. A detailed review of expectancies in double-blind RCTs is provided in Chapter 2. Briefly, there is preliminary evidence to suggest that participants’ beliefs about whether they have been allocated to receive active treatment or placebo affect their actual treatment responses (e.g. Bausell, Lao, Bergman, Lee, & Berman, 2005; Dar, Stronguin, & Etter, 2005). If so, this can greatly
reduce the validity of these types of trials, especially when participant blinding fails. A review of the possible contribution of the placebo effect to treatment side effects is provided in Chapter 5. In essence, evidence for negative placebo effects when participants or patients expect adverse outcomes, side effects being reported in placebo groups of double-blind RCTs, a relationship between expectancy and post-chemotherapy nausea, and some experimental studies manipulating the information participants receive about side effects, has raised the possibility that the placebo effect could cause or exacerbate some side effects (e.g. Barsky et al., 2002). If so, warning participants about treatment side effects could lead them to expect and, thereby, experience more side effects.

To date the evidence supporting the claims that the placebo effect influences participants’ outcomes in double-blind RCTs and contributes to treatment side effects is quite limited. The current project, therefore, aimed to investigate these two aspects of the placebo effect more thoroughly. In each case this was done in an applied clinical setting with patients and experimentally with healthy volunteers. To examine the impact of expectancies in double-blind RCTs, I reanalysed a trial of acamprosate and naltrexone for alcohol dependence based on whether the participants believed they received active treatment or placebo (Chapter 3). In recognition of the limitations to these types of reanalyses, I developed an experimental model aimed at testing how observable changes affect participants’ beliefs about their treatment allocation and whether these beliefs affect their actual treatment responses in two dummy (placebo only) double-blind RCTs for cognitive performance (Chapter 4). To determine whether the placebo effect contributes to treatment side effects, I developed an experimental model to test whether warnings about side effects increase the actual
occurrence, frequency, or severity of side effects using placebo treatment for sleep
difficulty in otherwise healthy volunteers (Chapter 6). This raised the possibility that
placebo-induced side effects exist, but are only small. To address this possibility, I
examined the extent to which patients’ expectancies for nausea were associated with
their post-chemotherapy nausea in a large sample of first time chemotherapy patients
(Chapter 7). Implications and limitations of these studies are discussed in Chapter 8.
Chapter 2: Review of Expectancies in Double-Blind RCTs

2.1 Double-Blind Randomised Placebo-Controlled Trials

Double-blind randomised placebo-controlled trials (RCTs) are generally considered the benchmark for establishing the efficacy of a treatment. They are almost always required by the U.S. Food and Drug Administration for new drug approval (Shapiro & Shapiro, 1997b). This is because they aim to determine the true efficacy of a treatment, above and beyond the placebo effect and other forms of bias.

As their name suggests, there are three integral components to double-blind RCTs: a) placebo control, b) randomisation, and c) blinding. Placebo control involves comparing the treatment of interest (active treatment) with a placebo rather than with no treatment. This means that all participants in the trial engage in the treatment process, but only those in the active treatment group receive the specific component of the active treatment that is being tested. Randomisation involves randomly assigning participants to receive either the active treatment or placebo in an attempt to distribute any possibly confounding participant characteristics across these two groups evenly. This avoids biases that might arise if, say, those most likely to benefit are allocated to the active treatment (Altman & Bland, 1999). Blinding involves keeping which group the participants have been allocated to hidden from those involved in the trial.

There are, in fact, three types of blinding possible in RCTs: single, double, and triple. Single blinding is when only the participants are unaware of their treatment allocation. In order to achieve this, participants are told that there is a chance that they
will receive active treatment or a placebo but that they will not know which one they have been given. Double blinding involves keeping the participants’ treatment allocation hidden from those administering the treatment as well as the participants. Triple blinding takes this one step further and ensures that researchers and/or assessors that have any contact with the participants are also unaware of the participants’ treatment allocation (Shapiro & Shapiro, 1997b). The three levels of blinding, then, can be thought of as; participant blinding, administrator blinding, and assessor blinding. Double-blind RCTs entail participant and administrator blinding, but not assessor blinding.

In much the same way that randomisation is employed to distribute potentially confounding participant characteristics across groups, participant blinding is aimed at distributing expectancies evenly across groups. The rationale for this is that, if expectancies are balanced evenly across groups, then the placebo effect in each arm should be equivalent and any differences between the active treatment and placebo should be a direct result of the active treatment alone (Nash, 1962). As such, some participants in the treatment group may believe they are receiving active treatment, others may think they are on placebo, and others still might be unsure. Similarly, some participants in the placebo group might believe they are on active treatment, others may believe they are on placebo, and others still might be unsure. Importantly, it is not that participants will have no expectancy regarding their treatment allocation but that these expectancies are as evenly distributed across groups as possible.

In addition to controlling for the placebo effect, blinding participants should also reduce sources of participant bias, such as demand characteristics and motivation.
to change, that may influence the trial outcome as these should also be evenly distributed across groups. In a similar fashion, administrator blinding is intended to eliminate experimenter bias and control the context in which the treatment is delivered. Although often unintentional, experimenter bias may lead to biased interpretation of symptoms/outcomes that favour a preconceived hypothesis about the treatment or to differences in communication towards participants receiving active treatment and those receiving placebo (see Rosenthal, 2002 for a review). By ensuring that administrators are unaware of the participants’ treatment allocation any such bias should be equivalent in both the active treatment and placebo groups. Taken together, placebo control, randomisation, and blinding should mean that differences between the active treatment and placebo are real and occur as a direct result of the treatment’s efficacy and not the placebo effect, participant bias, or experimenter bias. This is why double-blind RCTs are considered such a powerful method for assessing the true efficacy of a treatment.

2.2 How Expectancies Can Limit the Validity of Double-Blind RCTs

Despite employing placebo control and blinding there are at least three ways in which participant expectancies can limit the validity of double-blind RCTs. The first of these is when participant blinding is unsuccessful, meaning that expectancies have not been adequately controlled. This can reduce the internal validity of these types of trials because the effect of the active treatment cannot be dissociated from the placebo effect. The second involves the possibility that the placebo effect could mask some or all of the active treatment’s effect, making it difficult to determine the magnitude of the true treatment effect, which can also reduce the internal validity of the trial. The third is that informing participants that they might receive a placebo
could reduce the external validity of double-blind RCTs because in standard clinical practice there is, presumably, no doubt that an active treatment is being administered.

Before discussing each of these potential limitations it is important to distinguish between the two types of expectancies with which participants are faced in double-blind RCTs: they are 1) expectancies regarding treatment allocation and 2) expectancies regarding treatment efficacy. Expectancies regarding treatment allocation (perceived treatment) involve beliefs about whether the participant has been allocated to receive active treatment or placebo and can be characterised by statements such as; “I believe I am taking active treatment” or “I believe I have been given the placebo”. Expectancies regarding treatment efficacy, on the other hand, involve beliefs about whether the treatment is likely to be effective. They can be characterised by statements such as “I believe this treatment is effective and will improve my symptoms” or “I do not believe this treatment is effective, it will not affect my symptoms”. Although expectancies regarding treatment allocation and efficacy are distinct, they are very likely to interact to produce an overall expectancy regarding the outcome of the treatment delivered in the double-blind RCT. For example, a participant may have a strong belief that the active treatment is effective, but may believe that he/she has been allocated to receive a placebo, in which case, he/she would likely expect no improvement. Alternatively, another participant may believe that the active treatment is highly effective and believe he/she has been allocated to receive active treatment in which case, he/she would likely expect improvement. In this way, perceived treatment might be thought of as the initial level of expectancy in double-blind RCTs that can activate or deactivate expectancies regarding efficacy.
It is also important to note that both these types of expectancies will vary in strength. That is, one participant may have a very strong belief and feel certain that he/she has been allocated to receive active treatment, while another might have only a slight inclination that he/she is receiving active treatment. Similarly, a participant might believe very strongly that the active treatment is effective and expect large improvement, while another might believe that the treatment will have only a small effect. This makes the interaction between perceived treatment and expected efficacy more complicated because it is not necessarily clear what strength the perceived treatment must be to activate or deactivate the expected efficacy nor whether this relationship is additive. Nonetheless, if a participant does believe he/or she is on placebo, then this will likely diminish his/her overall expectancies regarding treatment outcome more than if he/she believed he/she was on active treatment regardless of strength. Expectancies for improvement are, therefore, likely to be stronger, on average, in those who believe they are receiving active treatment than those who believe they are on placebo.

2.2.1 Failed blinding

The first way in which expectancies can limit the validity of double-blind RCTs is when blinding fails. In these trials, accurate comparison between treatment and placebo is predicated on participants being unaware of their true treatment allocation so that expectancies do not bias outcomes in either arm. If blinding is not maintained, in that participants can guess their treatment allocation at a rate better than chance, then differences between the active treatment and placebo groups could occur as a result of the treatment alone, expectancy alone, or a combination of these two (Fergusson, Glass, Waring, & Shapiro, 2004; Fisher & Greenberg, 1993;
Swatzman & Burkell, 1998). This is because participants who know they are receiving active treatment are far more likely to expect improvement than participants who know they are receiving a placebo. Failed blinding, then, means that expectancies are not being adequately controlled and therefore presents a serious limitation to the internal validity of the trial. When blinding fails, the trial more closely resembles an open treatment versus no treatment comparison rather than the intended double-blind active treatment versus placebo comparison.

Despite this possibility, very few researchers conducting double-blind RCTs assess whether blinding has been maintained. Several recent reviews have examined the frequency of checking for the success of blinding in a range of different RCTs. Karanicolas et al. (2008) found that only 3 (2%) of 171 double-blind RCTs of orthopaedic trauma published between 1995 and 2004 reported assessing whether blinding had been maintained. Fergusson et al. (2004) took a random sample of 100 medical and 100 psychiatry double-blind RCTs published in selected high impact journals between 1998 and 2001 and found that only 15 (8%) of the 191 eligible trials reported that blinding had been assessed. These low rates are supported by researchers attempting to review common methods for assessing blinding finding only limited numbers of double-blind RCTs that have taken this precaution (e.g. Boutron, Estellat, & Ravaud, 2005; Fisher & Greenberg, 1993)

One criticism of these reviews might be that blinding was assessed but not reported in some of these trials. If so, these reviews may underestimate the true rate of testing for blinding. Hróbjartsson, Forfang, Haahr, Als-Nelisen, and Brorson (2007) investigated this possibility in their review of all double-blind RCTs indexed in The
Cochrane Central Register of Controlled Trials and published in 2001. They identified 1599 trials of which only 31 (2%) reported testing for the success or failure of blinding. They then took a random sample of 200 of the 1568 that did not report testing for blinding and sent questionnaires to the authors assessing whether blinding had been assessed but not reported. Of the 130 authors to respond, 15 (12%) did conduct some such test but failed to report the results. This suggests that assessing only published reports of testing for blinding does underestimate the true rate of testing, but only slightly. If Hróbjartsson et al.’s (2007) results are representative, then the number of double-blind RCTs that do test for blinding is around 14%, which is still very low considering the threat that failed blinding poses to the internal validity of these trials.

Perhaps even more alarming is that, when assessed, blinding is often found unsuccessful (e.g. Margraf et al., 1991; Morin et al., 1995; Rabkin et al., 1986). Estimates from Fergusson et al. (2004), Hróbjartsson et al. (2007), and Karanicolas et al.’s (2008) reviews suggest that unsuccessful blinding occurs in anywhere between 23%-60% of double-blind RCTs, with a high proportion of the remaining providing either insufficient or unclear data on the results of blinding. Many double-blind RCTs, therefore, cannot validly determine whether differences between active treatment and placebo result from the treatment itself or expectancy. Importantly, this is not to say that the active treatment is definitely ineffective in trials in which blinding is broken. In fact, participants’ ability to correctly guess their treatment allocation is often related to improvement in their symptoms (see Shapiro & Shapiro, 1997b for a review), which means that the probability of blinding being broken is actually likely to increase with the treatment’s efficacy (Sharpe, Ryan, Allard, & Sensky, 2003).
Nonetheless, failed blinding does mean that it is impossible to rule out expectancy as the cause of the active treatment’s observed superiority over placebo and this contradicts the aim of double-blind RCTs.

This concern is particularly relevant given increasing evidence that perceived treatment does impact on patient outcomes (e.g. Bausell et al., 2005; Dar et al., 2005; Lewis et al., 1975; McRae et al., 2004; Thomas et al., 2008). For example, in their double-blind RCT of dopamine transplantation and sham surgery for patients with Parkinson’s disease, McRae et al. (2004) found that perceived treatment predicted patient outcomes better than actual treatment allocation. Patients who believed they had received the transplantation demonstrated greater improvement in symptoms and reported higher quality of life than those who believed they had received the sham surgery, while there were only minimal differences between the dopamine transplantation and the sham surgery. Similarly, Bausell et al. (2005) observed no difference between real and sham acupuncture for pain following dental surgery in two single-blind RCTs, but found that those who believed they had received real acupuncture reported less pain than those who believed they had received sham acupuncture. These studies indicate that perceived treatment can be associated with participants’ treatment responses and reinforce the notion that failed blinding can invalidate double-blind RCTs.

Perhaps the most compelling evidence, however, comes from Dar et al.’s (2005) reanalysis of a double-blind RCT of nicotine replacement therapy for smoking cessation. The researchers conducting the original trial (Etter, Laszlo, Zellweger, Perrot, & Perneger, 2002) had observed that smokers receiving nicotine replacement
therapy smoked fewer cigarettes per day than those receiving placebo therapy and took this as evidence that nicotine replacement therapy is effective beyond the placebo effect. However, as with most double-blind RCTs, they had not considered participants’ perceived treatment beyond testing for the success of blinding at the end of the trial. When Dar et al. (2005) reanalysed the trial by including perceived treatment as a factor in addition to treatment allocation the difference between nicotine replacement therapy and placebo therapy disappeared. That is, the active treatment’s observed superiority over placebo was no longer evident when expectancies were controlled for. Furthermore, in both the active and placebo groups, participants who believed they were receiving nicotine smoked fewer cigarettes per day than those who believed they had been given a placebo, suggesting that perceived treatment was more important than actual treatment allocation.

While these findings demonstrate a strong association between perceived treatment and outcomes, there is some difficulty in interpreting the causal nature of this relationship. A placebo based interpretation would argue that participants who believe they are on active treatment will expect greater improvement than those who believe they are on placebo and that this will lead to greater improvement in the former. However, an alternative interpretation is that participants only come to believe they are taking active treatment because they experience improvement over the course of the trial. Here, improvement could result from the treatment itself or other factors, such as natural history. Thus, perceived treatment might influence treatment outcomes or treatment outcomes might influence perceived treatment.
The studies described above are insufficient to differentiate between these two accounts because they are correlational in nature. While participants are randomly allocated to receive active treatment or placebo there is no control over whether participants believe they are receiving active treatment or placebo. Instead, participants are simply asked to retrospectively rate whether they believe they were given active treatment or placebo at the end of the trial. Although some studies have asked participants to rate their perceived treatment at multiple time points and as early as one week into the treatment (Margraf et al., 1991; Morin et al., 1995), there is still the possibility that some improvement has occurred by this time, not to mention that repeatedly questioning participants about their perceived treatment might affect their responses.

Even if the possible influence of treatment outcomes on perceived treatment could be ruled out, there are other factors that could account for perceived treatment’s effect on treatment outcomes. For example, those who believe they have been allocated to active treatment might be more compliant than those who believe they have been given placebo or might be more motivated to improve and engage in other beneficial activities.

While these concerns make it impossible to determine whether perceived treatment affects treatment outcomes, they do not eradicate the threat failed blinding poses to the internal validity of double-blind RCTs. The intention of double-blind RCTs is to control for the placebo effect and deduce whether any observed improvement results from the active treatment or expectancy. When blinding fails and participants can guess their allocation at a rate better than chance, expectancies are not
being properly controlled and this means that the researchers cannot be confident about the cause of any observed improvement in the treatment group.

2.2.2 Expectancies masking the active treatment’s effect

The second way in which expectancies can limit the validity of double-blind RCTs concerns the possibility that the placebo effect could mask the active treatment’s effect. As discussed above, in double-blind RCTs the difference between the active treatment and placebo is believed to reflect the active treatment’s true efficacy beyond the effects of expectancy and other forms of bias. So, when blinding is maintained and those receiving active treatment show greater improvement than those receiving placebo, the active treatment is taken to be efficacious over and above any placebo effect. Conversely, when there is no difference between those receiving active treatment and placebo any improvement in those receiving active treatment is attributed to the placebo effect.

The central assumption on which this is based is that the treatment effect and placebo effect are additive. However, while treatment and placebo responses are often discussed as though the placebo effect simply adds onto the treatment effect, there are neither empirical nor theoretical reasons to support this view (Kirsch, 2000; Kirsch, Moore, Scoboria, & Nicholls, 2002). For example, Juliano and Brandon (2002) induced anxiety in briefly abstinent smokers and gave them cigarettes with either nicotine or no nicotine but told half of them that the cigarettes contained nicotine and the other half that the cigarettes were placebos. In those who received the placebo, smoking urges were lower for those who were told they would be receiving nicotine, indicating a placebo effect. In those given nicotine, however, there was an equal
reduction in smoking urges for those told to expect nicotine and those told to expect placebo. That is, nicotine reduced smoking urges without expectancy, but the addition of expectancy did not lead to further reductions despite evidence for a placebo effect. Importantly, this finding held when participants who did not believe the expectancy manipulation were excluded from the analysis. Evidence that placebo responses can differ to responses to hidden treatment (e.g. Hull & Bond, 1986) provides further support to the idea that active treatment and placebo effects may not be additive. A potential lack of additivity between treatment and placebo effects is perhaps unsurprising given that in classical conditioning some conditioned responses differ from the unconditioned response (e.g. Black, 1971; Glaudier, Drummond, & Remington, 1992; Siegel, 1975).

If treatment and placebo responses are not additive, then the difference between the active treatment and the placebo will not accurately reflect the treatment’s effect without expectancy. This is particularly relevant when there is equal improvement in both the active treatment and placebo groups. In this instance, the most common conclusion would be that the active treatment only produced improvement via the placebo effect. An alternative explanation of this, however, is that the active treatment is efficacious without expectancy, but there were strong placebo effects in both the active and placebo groups that masked the active treatment’s effect. This possibility is a product of the fact that participant blinding attempts to evenly distribute expectancies regarding perceived treatment across the active treatment and placebo groups, as does randomisation for expectancies regarding treatment efficacy, rather than eradicating them. As such, participants will
still have expectancies regarding the trial outcome and this means that placebo effects within each arm are possible.

Linde et al.’s (2007) review of four double-blind RCTs comparing real acupuncture, placebo acupuncture, and a waitlist control for pain supports this possibility. While they found no differences between real acupuncture and placebo acupuncture, patients who had higher expectancies regarding the efficacy of acupuncture benefited more from the treatment than those who had low expectancies regarding acupuncture’s efficacy, regardless of what they actually received. Although they did not test for this possibility, Linde et al. (2007) argued that strong expectancies for improvement could create a ceiling effect whereby it is impossible to detect the effect of the treatment alone, i.e. without expectancy. If this is the case, then a lack of difference between the active treatment and placebo may not indicate that the active treatment has no unique effect. The same treatment delivered without participants’ awareness might lead to improvement, as with Juliano and Brandon’s (2002) study involving nicotine and denicotinised cigarettes administered with and without expectancy. This possibility limits the internal validity of double-blind RCTs because these trials cannot reliably determine whether the active treatment or the placebo effect has caused the improvement when there appears to be no difference between that active treatment and placebo groups.

2.2.3 Differences between standard clinical practice and double-blind RCTs

The difference in expectancies that participants in double-blind RCTs are faced with compared with standard clinical practice reflects the third way in which double-blind RCTs can be limited. In double-blind RCTs participants will validly
question whether they have been allocated to receive active medication because they know there is a chance that they could receive a placebo. In standard clinical practice, however, it is highly unlikely that patients would doubt that they are receiving active medication (Barfod, 2005; Kirsch & Weixel, 1988; Nash, Holroyd, Rokicki, Kvaal, & Penzien, 2002). That is, while expectancies regarding the efficacy of the treatment can affect both participants in double-blind RCTs and patients receiving treatment in standard clinical practice, questions regarding perceived treatment are specific to participants in double-blind RCTs and other blinded placebo-controlled trials. If this translates into differences in overall expectancies regarding the outcome of the treatment, then one might expect greater efficacy for the same treatment in standard clinical practice, where there is no doubt that the treatment is active, compared with double-blind RCTs, where some doubt is likely, provided that the condition being treated is placebo-responsive. Outcomes found in double-blind RCTs may not, then, accurately reflect outcomes that would be found in standard clinical.

Relatively few studies have examined whether treatment responses differ under double-blind RCT administration compared with standard clinical practice. Skovlund (1991) compared the analgesic effect of paracetamol for uterine pain in postpartum women across two consecutive trials. The first trial was a double-blind RCT in which paracetamol was tested against placebo and, as such, women were aware that there was a possibility that they would receive a placebo. The second trial was also double-blind, but it was not placebo-controlled. Instead, paracetamol was tested against naproxen meaning that women in this trial, although unsure of the type, knew that they would be receiving an active analgesic, which approximates standard clinical practice. Women in the second trial reported lower uterine pain than those in
the first trial despite receiving the same dose of paracetamol. This supports the notion that treatment responses in double-blind RCTs can underestimate responses to the same treatment administered in standard clinical practice.

Rochon et al. (1999) compared receipt of nonsteroidal anti-inflammatory drugs for rheumatoid arthritis in 25 double-blind RCTs with 33 double-blind trials comparing the same treatment to another active treatment. As did Skovlund (1991), they found greater efficacy for the same treatment when participants knew they would be receiving an active treatment compared with when they knew there was a chance they were receiving a placebo. One limitation to both these studies, however, is that participants were not randomly allocated to the two types of treatment administration and this may have biased results.

Pollo et al. (2001) and Kirsch and Weixel (1988) conducted similar studies but used placebo treatment rather than active treatment and properly randomised their participants. Pollo et al. (2001) gave cancer patients recovering from thoracic surgery a continuous saline infusion. One group of patients was told that the infusion was a powerful painkiller, reflecting instructions analogous to standard clinical practice. A second group was told that the infusion might contain a powerful painkiller or saline, reflecting double-blind RCT administration. A third, control group was told that the infusion contained saline. Over the next three days, the patients who were told they were receiving a powerful painkiller required fewer analgesics compared with both those who were told they may receive a placebo and the control group, despite identical pain ratings. This implies a greater placebo effect in conditions analogous to
standard clinical practice compared with double-blind RCT administration, because fewer analgesics were required to achieve the same level of pain reduction.

Kirsch and Weixel (1988) administered decaffeinated coffee to their participants but told half that they were receiving caffeinated coffee, simulating instructions given in standard clinical practice, and told the other half that they were receiving either caffeinated or decaffeinated coffee, simulating double-blind RCT administration. As with Pollo et al. (2001) they found a larger placebo effect under standard clinical practice administration than double-blind RCT administration, in that mean heart rate was higher for those who were told they were receiving caffeine than those who were told they would receive either caffeine or placebo.

Another small group of studies has examined whether there is an interaction between the instructions provided to participants and the treatment they receive, i.e. active or placebo, with less consistent results both within and across studies (Hughes, Gulliver, Amori, Mireault, & Fenwick, 1989; Kirsch & Rosadino, 1993; Nash et al., 2002). For example, while Kirsch and Rosadino (1993) found that caffeinated coffee increased tension more than decaffeinated coffee only when participants were told they would be receiving caffeine and not when they were told they may receive caffeinated or decaffeinated coffee, they also found that caffeinated coffee increased alertness more than decaffeinated coffee regardless of whether participants were told they would be receiving caffeine or that they might receive caffeine or a placebo.

These inconsistent results are particularly difficult to disentangle given the general paucity of research in this area. What does seem clear, however, is that at least
under some circumstances double-blind RCT administration produces different
treatment responses to standard clinical practice. Further, when this does occur, it
seems that the most likely result is that double-blind RCT administration
underestimates the magnitude of the treatment response that would be obtained in
standard clinical practice (Kirsch & Rosadino, 1993; Kirsch & Weixel, 1988; Nash et
al., 2002; Pollo et al., 2001; Skovlund, 1991 with only one exception, Hughes et al.,
1989).

While it is tempting to attribute any differences between double-blind RCTs
and standard clinical practice to participant expectancies, there is another possible
explanation. As mentioned earlier, the intention of participant blinding is to control
for participant bias, such as demand characteristics and motivation to change, not just
participant expectancies. This means that manipulating the instructions that
participants receive about their treatment could also affect participant bias. For
example, participants told that they have been given an active medication might feel
more compelled to report improvement than participants who know there is a chance
that they are receiving a placebo. This may seem like a particularly relevant
possibility given that the majority of the above findings were for subjective outcomes.
However, the higher heart rate in response to decaffeinated coffee for those who were
told they were receiving caffeine compared with those who were told they would
receive either caffeine or placebo observed by Kirsch and Weixel (1988) seems
unlikely to have resulted from participant bias and points towards expectancy as a
causal mechanism. There is also evidence from other studies that instructional
expectancy manipulations can have a powerful effect on objective treatment responses
(e.g. Benedetti, Maggi et al., 2003; Flaten, Simonsen, & Olsen, 1999). So it remains
likely that the differences in expectancies created in double-blind RCTs and standard clinical practice can cause differences in responses to the same treatment via the placebo effect in at least some circumstances.

There is, however, some caution required in interpreting the extent to which this limits double-blind RCTs. The primary aim of these trials is to determine the efficacy of an active treatment without expectancy, so the extent to which they replicate standard clinical practice is probably less important. It does, on the other, demonstrate that researchers and clinicians cannot reliably predict outcomes that would be achieved in standard clinical practice from double-blind RCTs. Further, it supports the notion that perceived treatment can impact on outcomes in these trials and that considering perceived treatment is essential for fully understanding treatment responses in double-blind RCTs.

2.3 Alternatives to double-blind RCTs

In recognition of these limitations a number of alternatives to standard double-blind RCTs have been suggested. These are 1) the balanced placebo design, 2) the ‘open versus hidden’ design, and 3) the use of active rather than benign placebos as the control comparison. While each of these alternatives overcomes at least one problem associated with double-blind RCTs, they are not faultless. In most cases, overcoming one problem is associated with the introduction of another. Of course, some clinical trials avoid using placebo controls by comparing two active treatments, usually a standard treatment with established efficacy versus some new treatment. However, this approach is not considered as a true alternative to double-blind RCTs here, because these trials do not aim to determine the efficacy of a treatment above and beyond the placebo effect, they simply attempt to establish relative efficacy.
2.3.1 The Balanced Placebo Design

The balanced placebo design was developed by Ross, Krugman, Lyerly, and Clyde (1962). It consists of a 2x2 factorial design with treatment allocation and expected allocation as factors. Participants are allocated to receive the active treatment or a placebo and are either told that they have been given the active treatment or that they have been given a placebo. This produces four conditions as shown in Figure 2.1. One group (cell A) is given active treatment and are aware of this and, as such, should demonstrate the combined effects of the active treatment and the placebo effect. A second group (cell B) also receives active treatment but is told that they have been given a placebo. Any response to the treatment in this group should result from the active treatment alone. A third group (cell C) is given a placebo but is told they have been given the active treatment which means that any responses observed in this group should be the result of the placebo effect. Finally, a fourth group (cell D) receives a placebo and is aware of this. This group can, therefore, be considered a natural history group, in that no active treatment or placebo effects should occur.

**PARTICIPANTS TOLD TO EXPECT:**

<table>
<thead>
<tr>
<th></th>
<th>Active</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active with expectancy</td>
<td>Active treatment with expectancy</td>
<td>Placebo with expectancy</td>
</tr>
<tr>
<td>(A)</td>
<td>(B)</td>
<td>(C)</td>
</tr>
<tr>
<td>Placebo without expectancy</td>
<td></td>
<td>Placebo without expectancy</td>
</tr>
<tr>
<td>(D)</td>
<td>(D)</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 2.1. The balanced placebo design: participants are allocated to receive either active treatment or placebo with half being told that they will receive active treatment and the others being told that they will receive a placebo.*
A number of studies investigating the effects of alcohol (see Hull & Bond, 1986; Rohsenow & Marlatt, 1981 for reviews) and more recently nicotine (e.g. Gottlieb et al., 1987; Juliano & Brandon, 2002; Kelemen & Kaighobadi, 2007; Perkins et al., 2004) have employed the balanced placebo design. The main benefit of this method is that it can isolate the proportion of the treatment response that is uniquely attributable to the active treatment, the proportion of the treatment response that is uniquely attributable to the placebo effect, and, importantly, whether these two are additive (Kirsch et al., 2002). In this regard, the balanced placebo design is superior to double-blind RCTs because, without proof of additivity, the only conclusions the latter can validly draw is that the active treatment is more efficacious than placebo under double-blind RCT conditions.

The main problem with the balanced placebo design is that the deliberate deception inherent in this procedure is generally considered unethical (Kirsch et al., 2002; Miller, Wendler, & Swartzman, 2005; Swatzman & Burkell, 1998; Waring, 2008). While there has been ongoing debate regarding whether administering a placebo in double-blind RCTs is ethical (Colloca, Lopiano, Lanotte, & Benedetti, 2004; Shapiro & Shapiro, 1997b), there is no deception involved in these trials. Participants are informed that they will either receive active treatment or a placebo and this is exactly what happens (Miller et al., 2005; Waring, 2008). In the balanced placebo design, some participants are told they are receiving an active treatment when in fact they have been given a placebo and others are told they have been given a placebo, but are actually given active treatment. This makes double-blind RCTs much more ethically favourable than the balanced placebo design and is the most likely
reason that the balanced placebo design has only been used in healthy participants to date (Miller et al., 2005).

2.3.2 The ‘Open versus Hidden’ Design

The ‘open versus hidden’ design involves comparing the efficacy of an active treatment in a group of participants who are aware that the treatment is being administered (open treatment), with a group of participants who are unaware that the treatment is being delivered (hidden treatment). The response to the hidden treatment is taken as the efficacy of the active treatment without expectancy and the response to the open treatment is taken as the active treatment plus any associated placebo effect. The main benefit of the ‘open versus hidden’ design, then, is that it can isolate the component of a treatment response that is attributable to the active treatment alone, which is not possible in double-blind RCTs. A second benefit, is that all participants receive active treatment, making it a better ethical alternative to both double-blind RCTs and the balanced placebo design (Benedetti, Maggi et al., 2003; Colloca et al., 2004; Finniss & Benedetti, 2005; Kirsch, 2003).

Despite their greater ethical acceptability and usefulness for determining differences between an active treatment with and without expectancy, the ‘open versus hidden’ design is possible in only very limited circumstances (Colloca et al., 2004). Most of the studies employing this design have been conducted in hospitals and involved intravenous injections delivered mechanically, a treatment context which is amenable to hidden administration (e.g. Amanzio, Pollo, Maggi, & Benedetti, 2001; Benedetti, Maggi et al., 2003; Gracely, Dubner, Wolskee, & Deeter, 1983; Levine & Gordon, 1984). However, the majority of treatment will not be amenable. For example, it seems impossible to engineer hidden administration of oral
pills, a topical cream, or surgery. Further, comparing open and hidden treatment using this design does not control for the possibly confounding effects of participant bias. Participants receiving open treatment may be more subject to demand characteristics or more motivated to change than those receiving hidden treatment. Thus, not only is the ‘open versus hidden’ design impractical for the majority of treatments, it also re-introduces the problem of participant bias.

2.3.3 Active rather than benign placebos

The final alternative research design involves a slight modification to the standard double-blind RCT. It involves using an active placebo, which mimics the side effects of the active treatment, rather than a benign placebo as the control comparison. For example, atropine has been used as an active placebo for trials of tricyclic antidepressants (Gaudiano & Herbert, 2005). If participants’ guesses regarding their treatment allocation are influenced by side effects, then active placebos should make blinding harder to break because side effects should be equivalent in both groups (Edward, Stevens, Braunholtz, Lilford, & Swift, 2005; Fisher & Greenberg, 1993; Gaudiano & Herbert, 2005; Quitkin, 1999). As a result, active placebos may increase the internal validity of double-blind RCTs.

Although guesses about perceived treatment are not always based on side effects (e.g. Morin et al., 1995; Rabkin et al., 1986), in at least some cases they are (e.g. Margraf et al., 1991; Moscucci, Byrne, Weintraub, & Cox, 1987; or see Shapiro & Shapiro, 1997b for a review), which suggests that using active placebos may improve the success of blinding. A least one study has shown improved blinding as a result of using an active placebo (Turner, Jensen, Warms, & Cardenas, 2002). Further,
double-blind RCTs using active placebos tend to show diminished antidepressant effects compared with trials using benign placebos (Gaudiano & Herbert, 2005), which suggests that active placebos might ensure a more balanced distribution of expectancies across treatment arms.

Yet there are two important problems with employing active placebos: ethicality and practicality. Active placebos are used to emulate the adverse side effects of the active treatment and thereby produce adverse effects in those allocated to receive them. This means that participants in the active placebo group are likely to suffer at least some discomfort from their involvement in the trial in addition to being denied a potentially useful treatment (Boutron et al., 2006; Gaudiano & Herbert, 2005; Max, 2007). As such, careful consideration must be given regarding the benefits and costs of administering an active placebo, which, in some circumstances, might undermine the ethicality of using these as controls in double-blind RCTs (Edward et al., 2005). The second criticism is that it might not always be easy to identify an appropriate active placebo and to be certain that its active component has no effect on the condition of interest (Max, 2007). If the active placebo does influence the condition being studied, whether positively or negatively, then the efficacy of the active treatment without expectancy cannot be validly determined.

In addition to these problems, double-blind RCTs that employ active placebos rather than benign placebos can still be limited because of the possibility that the placebo effect masks the active treatment and as a result of the difference in instructions between them and standard clinical practice. Further, although Turner et al. (2002) argued that their use of an active placebo improved blinding compared with
other trials employing benign placebos, 70% of their participants were still able to
correctly guess that they had been allocated to receive active treatment. Thus, using
active placebos does not necessarily overcome any of limitations introduced to
double-blind RCTs as a result of participant expectancies.

2.4 Conclusions

Double-blind RCTs do not eradicate expectancies, but simply aim to control
for them. Participants in these trials will form expectancies both about the efficacy of
the treatment (expected efficacy) and whether or not they have been allocated to
receive active treatment or placebo (perceived treatment). There are a number of ways
that these expectancies can undermine the validity of double-blind RCTs. Firstly,
researchers conducting these trials rarely assess whether blinding has been maintained
and, when they do, it is often found to be unsuccessful. This means that differences in
perceived treatment could contribute to any observed differences between the active
treatment and placebo groups and this thereby limits the internal validity of the trial.
Secondly, double-blind RCTs do not test for the additivity of the active treatment’s
effect and the placebo effect, which means that a failure to detect a difference
between the active treatment and the placebo is not sufficient evidence that the active
treatment is ineffective without expectancy. Thirdly, expectancies regarding
perceived treatment are specific to double-blind RCTs and other trials that employ
participant blinding, which means that double-blind RCTs may underestimate
treatment outcomes that would be obtained in standard clinical practice.

By far the biggest limitation to understanding the full impact of expectancies
in double-blind RCTs is that the research to date cannot determine whether perceived
treatment has a causal impact on actual treatment responses. The main reason for this is that there are relatively few studies investigating the role of perceived treatment in double-blind RCTs and, in those that have, there is very little attempt to control for other possible contributing factors, such as demand characteristics and motivation. Thus, while perceived treatment is often found to be related to treatment responses, it is unclear whether participants who believe they are taking active treatment experience improvement as a result of that expectation or whether experiencing improvement causes a participant to believe that he/she is taking active treatment. Equally possible is that any impact of perceived treatment on actual treatment responses is mediated by motivation or other similar psychological variables. For example, a person who believes he/she is taking active treatment may be more motivated to improve than a person who believes he/she is taking placebo and this increased motivation may result in greater improvement, not the expectancy.

In an attempt to overcome these limitations, I examined whether perceived treatment predicted actual treatment responses in a previous double-blind RCT of alcohol dependence in which blinding was maintained and after controlling for motivation to change (Chapter 3). In two experimental studies I also tested whether observable improvement affects perceived treatment and whether this, in turn, affects treatment responses. This involved conducting dummy (placebo only) double-blind RCTs and providing participants with false feedback about their performance in order to experimentally manipulate their perceived treatment (Chapter 4).
Chapter 3: Perceived Treatment in a Double-Blind RCT for Alcohol Dependence (Study 1)

3.1 Introduction

This study investigated the relationship between perceived treatment and treatment outcomes in a double-blind RCT comparing acamprosate, naltrexone, and placebo for alcohol dependence. This consisted of reanalysing the trial as a 2x2 design with actual treatment and perceived treatment as factors. Of interest was whether perceived treatment was more reliably associated with the various indices of alcohol dependence and the number of adverse side effects than was actual treatment. As discussed in the previous chapter, this type of analysis has previously been conducted for double-blind RCTs of acupuncture for pain (Bausell et al., 2005), neurosurgery for Parkinson’s disease (McRae et al., 2004), smoking cessation (Dar et al., 2005; Thomas et al., 2008), and vitamin C for the common cold (Lewis et al., 1975). In all of these studies, participants who believed they were receiving active treatment showed greater improvement than participants who believed they were receiving a placebo, irrespective of the actual treatment they received.

The double-blind RCT reanalysed here was originally reported in Morley et al. (2006) and had a number of characteristics that made it particularly useful for exploring the relationship between perceived treatment and actual treatment responses. First, the researchers conducting the trial assessed participants’ perceived treatment at the end of the trial in order to test for the success for blinding, which made the current reanalysis possible. Second, blinding was maintained and there was no difference between the active treatments and placebo in the primary analysis. This
means that actual and perceived treatment should be relatively independent of each other compared with trials in which blinding fails. In all but one of the reanalyses mentioned above blinding was unsuccessful and that study (McRae et al., 2004) contained a very small sample who appeared overly optimistic about their chances of receiving the active treatment. Finally, the researchers assessed participants’ motivation to change before and after treatment which meant that this could be included as a covariate in the analysis. This was beneficial because perceived treatment could result in differences in motivation to change if, say, participants who believe they have been given a placebo become discouraged and feel that their participation in the trial will not benefit them. If this were the case, then poorer outcomes in these participants compared with those who believe they are taking active treatment could result from the former’s reduced motivation, rather than differences in expectancy.

3.2 Methods

A full description of the methods used in the original trial can be found in Morley et al. (2006). A summary of these is given below, along with the design and data analysis for the current analysis. It should be noted that, despite assessing perceived treatment, the researchers conducting the original trial used this information only to test whether blinding had been achieved. As such, the analysis reported below is entirely novel with respect to the original study.

3.2.1 Participants

One hundred and sixty-nine men and women diagnosed with alcohol dependence or abuse, according to the Diagnostic and Statistical Manual IV, took part
in the trial. They were recruited during attendance at an inpatient detoxification programme, an out-patient treatment follow-up, and via live and print advertisements.

3.2.2 Design

The reanalysis was based on a 2x2 design with actual treatment and perceived treatment as factors. In order to achieve this, the four treatment types, naltrexone, acamprosate, naltrexone placebo, and acamprosate placebo were simplified into either active treatment, i.e. acamprosate and naltrexone, or placebo. This was justified on the basis that there were no differences found between acamprosate and naltrexone in the primary analysis of the original trial. Perceived treatment was based on participants’ guesses regarding their treatment allocation at the end of the trial, which could be either active or placebo. The dependent variables were days until first relapse, total number of days abstinent, total alcohol consumption during the trial, posttreatment alcohol dependency, posttreatment alcohol cravings, and number of adverse side effects.

3.2.3 Measures

Drinks diary: Participants were required to complete a diary comprised of daily cards which assessed their alcohol consumption, compliance, and adverse side effects for each day they received treatment. Chick, Howlett, Morgan, and Ritson (2000), who were the first to use these daily cards, found a high level of consistency between them and serum γ-glutamyl transferase activity, an objective measure of alcohol consumption, suggesting they have good validity. The diary was used to derive the number of days until first relapse, the total number of days abstinent, total alcohol consumption during the treatment period, and number of adverse side effects.
Alcohol dependence: was assessed via the Alcohol Dependence Scale (ADS: Skinner & Allen, 1982). The scale contains 25 items that assess alcohol withdrawal symptoms, impaired control over drinking, awareness of compulsion to drink, increased tolerance to alcohol, and strength of drink-seeking behaviour. Higher scores on the ADS correspond to greater alcohol dependence.

Alcohol cravings: were assessed using the Penn Alcohol Craving Scale (PACS: Flannery, Volpicelli, & Pettinati, 1999). The scale contains 5 items that assess the frequency, intensity, and duration of thoughts about drinking. Higher scores on the PACS indicate more alcohol cravings.

Motivation to change: was assessed with the Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES: Miller & Tonigan, 1996). The scale was designed specifically for problem drinkers and contains 19 items that assess taking steps to overcome problem drinking, recognition of problem drinking, and ambivalence towards problem drinking. Higher scores on the SOCRATES correspond to greater motivation to change.

3.2.4 Procedure

After informed consent was obtained, participants were required to remain abstinent from alcohol for a minimum of 3 days, but not more than 21 days, prior to the baseline assessment, which involved assessing demographics, alcohol consumption, alcohol dependence and cravings, and motivation to change. Participants were then randomised to receive naltrexone, acamprosate, naltrexone
placebo or acamprosate placebo for 12 weeks and occasional compliance therapy sessions. The compliance therapy involved discussing problems that might affect treatment compliance, such as, ambivalence and misconceptions about medications (Teesson et al., 2003). During the treatment period participants were asked to keep a drinks diary which was used to monitor their alcohol consumption and cravings. At the end of the 12 weeks, participants underwent a posttreatment assessment of alcohol dependence and cravings, motivation to change, compliance, and perceived treatment. The study was approved by the Human Ethics Review Committees of Central Sydney Area Health Service, South Eastern Sydney Area Health Service and Wentworth Area Heath Service.

3.2.5 Statistical Analysis

Analysis of variance (ANOVA) and Fisher’s exact test were used to assess whether there were differences in baseline characteristics across actual treatment and perceived treatment. These included age, sex, abstinence and average number of drinks per day before the trial, and pretreatment alcohol dependence, alcohol cravings and motivation to change. Analysis of co-variance (ANCOVA) was used to test for differences in compliance, the number of counselling sessions attended, and posttreatment motivation to change at the end of the trial, while controlling for age, sex, and pretreatment abstinence, average number of drinks and alcohol dependence. Blinding was confirmed via a Chi-square test of independence.

ANCOVA was used to test the impact of actual treatment, perceived treatment, and their interaction on days until relapse, total days abstinent, total alcohol consumption, alcohol dependence, and alcohol cravings. Age, sex, baseline
abstinence, alcohol dependence and average number of drinks per day, number of pills involved in the treatment regimen, and posttreatment motivation to change were included as covariates. Alcohol cravings at baseline were also included as a covariate when posttreatment alcohol cravings was the outcome variable. ANCOVA also tested the impact of actual and perceived treatment on the number of adverse side effects. Here, age, sex, and number of pills involved in the treatment regimen were included as covariates. All statistical analyses were conducted using SPSS software (version 15; SPSS Inc, Chicago, Ill) and results were considered significant when \( p < .05 \).

### 3.3 Results

#### 3.3.1 Baseline characteristics

Of the 169 participants who were recruited, 116 (69%) completed the trial including responding to the question regarding their perceived treatment. They had a mean age of 45.7 years (range 23-66) and the majority were male (66%). Table 3.1 provides details of the baseline characteristics of these participants according to their actual treatment and perceived treatment. The only significant difference was for age,

| Table 3.1. Baseline characteristics across actual treatment and perceived treatment. Means (SE) are given for continuous variables, percentage male is given for sex, and * denotes \( p < .05 \). |
|----------------|----------------|----------------|----------------|----------------|
|                | Receive Active |                | Receive Placebo |                |
|                | Believed Active| Believed Placebo| Believed Active | Believed Placebo|
| Age*           | 47.5 (1.3)     | 47.3 (1.8)     | 44.4 (1.9)     | 41.7 (1.7)     |
| Sex (% male)   | 67.4           | 76.9           | 52.4           | 65.7           |
| Abstinence     | 4.7 (.53)      | 5.2 (.71)      | 5.0 (.79)      | 4.6 (.80)      |
| Drinks/day     | 13.0 (1.1)     | 11.7 (1.4)     | 10.2 (1.6)     | 13.0 (1.5)     |
| ADS            | 20.5 (1.3)     | 19.5 (1.7)     | 19.4 (1.9)     | 21.6 (1.8)     |
| PACS           | 18.7 (1.1)     | 18.3 (1.4)     | 18.1 (1.6)     | 15.9 (1.5)     |
| SOCRATES       | 27.4 (.48)     | 25.8 (.63)     | 26.6 (.70)     | 27.0 (.67)     |
which indicated that those receiving active treatment were significantly older than those receiving placebo treatment, $F(1,112)=6.32, p=.01$.

3.3.2 The success of blinding

Table 3.2 shows the distribution of participants according to their actual treatment and perceived treatment. A Chi-square test of independence revealed that there was no significant association between actual treatment and perceived treatment indicating that blinding was maintained, $\chi^2(df=1, n=116)=2.9, p=.07$. There was, however, a tendency for those receiving active treatment to be more likely to believe they had been given active treatment than placebo.

Table 3.2. Distribution of participants according to their actual treatment allocation and their perceived treatment allocation.

<table>
<thead>
<tr>
<th>Perceived Treatment</th>
<th>Active</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>46</td>
<td>26</td>
</tr>
<tr>
<td>Placebo</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>67</td>
<td>49</td>
</tr>
</tbody>
</table>

3.3.3 Total alcohol consumption

After controlling for baseline characteristics, number of pills, and motivation to change, the main effect of expectancy for total alcohol consumption was significant. As shown in Figure 3.1, participants who believed they had been allocated to receive active treatment reported consuming 147 less drinks than participants who believed they had been allocated to receive placebo, $F(1,89)=7.84, p<.01$. There was, however, no main effect of actual treatment and no interaction between this and perceived treatment, highest $F(1,89)=1.65, p=.20$. 

54
Figure 3.1. Covariate adjusted mean (+SE) total alcohol consumption by actual and perceived treatment. Those who believed they were receiving active treatment consumed significantly fewer drinks than those who believed they were receiving placebo. The main effect of actual treatment and the interaction were not significant.

Figure 3.2. Covariate adjusted mean (+SE) alcohol dependence score (ADS) across actual and perceived treatment. Those who believed they were taking active treatment reported significantly lower alcohol dependence than those who believed they were receiving placebo. No other differences were significant.
3.3.4 Alcohol dependence ratings

Alcohol dependence ratings at the end of the trial are presented in Figure 3.2. After controlling for baseline characteristics, number of pills, and motivation to change, the only significant difference was for perceived treatment. Participants who believed they received active treatment rated their alcohol dependence as 4.2 points lower than those who believed they had received the placebo, $F(1,89) = 4.48, p = .04$. There was no significant effect of actual treatment on alcohol dependence ratings, nor a significant interaction, both $F < 1$.

3.3.5 Alcohol cravings

A similar pattern of results was found for alcohol cravings, as shown in Figure 3.3. Participants who believed they had been given the active treatment reported alcohol cravings 5.6 points lower than participants who believed they had been given the placebo, after controlling for baseline characteristics, number of pills, and motivation to change, $F(1,87) = 8.77, p = .01$. Again, there were no significant differences for actual treatment and no interaction, highest $F(1,87) = 1.28, p = .26$.

![Figure 3.3](image_url)

*Figure 3.3. Covariate adjusted mean (+SE) alcohol cravings (PACS). Those who believed they were on active treatment reported significantly lower cravings than those who believed they were on placebo. No other differences were significant.*
3.3.6 Total number of days abstinent

Total number of days abstinent across actual and perceived treatment is shown in Figure 3.4. After controlling for baseline characteristics, number of pills, and motivation to change, both actual and perceived allocation were non-significant, both $F<1$. There was, however, a significant interaction which indicated that the difference in abstinence between expecting active treatment and expecting placebo was not the same for those receiving active treatment compared with those receiving placebo, $F(1,89)=4.50, p=.04$.

![Total Number of Days Abstinent](image)

Figure 3.4. Covariate adjusted mean (+SE) number of days abstinent across actual and perceived treatment. There were no main effects of actual and perceived treatment, although their interaction was significant.

3.3.7 Number of days until relapse

The mean number of days until relapse across actual and perceived treatment is shown in Figure 3.5. While there were tendencies towards longer time until relapse for those who believed they had received active treatment compared with those who believed they had been given placebo and for those actually receiving active treatment
compared with those receiving placebo, these differences were not significant after controlling for baseline characteristics, number of pills, and motivation to change, highest \( F(1,89)=1.91, p=.17 \). There was also no interaction between perceived treatment and actual treatment for number of days until relapse, \( F<1 \).

**Days Until Relapse**

![Graph showing Days Until Relapse](image)

*Figure 3.5. Covariate adjusted mean (+SE) number of days until relapse across actual and perceived treatment allocation. No differences were significant.*

3.3.8 Number of adverse side effects

Number of adverse side effects showed the same pattern as total alcohol consumption, alcohol dependence, and alcohol cravings. As shown in Figure 3.6, after controlling for baseline characteristics, number of pills, and motivation to change, participants who believed they received active treatment reported significantly more adverse side effects than participants who believed they received a placebo, MD=.48, \( F(1,109)=4.00, p<.05 \). There were no differences in the number of adverse side effects reported between those who received active treatment and those who received placebo, nor was there a significant interaction, highest \( F(1,109)=2.32, p=.13 \).
Number of Adverse Side Effects

<table>
<thead>
<tr>
<th>Believed Active</th>
<th>Believed Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Received Active</td>
<td>Received Placebo</td>
</tr>
</tbody>
</table>

Figure 3.6. Covariate adjusted mean (+SE) number of adverse side effects across actual and perceived treatment allocation. Participants who believed they received active treatment reported significantly more adverse side effects than those who believed they received placebo. No other differences were significant.

3.3.9 Compliance, counselling, and posttreatment motivation to change

Means for number of days compliant, number of counselling sessions attended, and posttreatment motivation to change are presented in Table 3.3. There were no significant differences in any of these variables for actual treatment, perceived treatment, and their interaction after controlling for age, sex, abstinence and average number of drinks per day before the trial, and pretreatment alcohol.

Table 3.3. Mean (SE) number of days compliant with treatment instructions (compliance), number of counselling sessions attended (counselling), and posttreatment motivation to change (SOCRATES) by actual and perceived treatment. No differences were significant.

<table>
<thead>
<tr>
<th></th>
<th>Receive Active</th>
<th></th>
<th>Receive Placebo</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expect Active</td>
<td>Expect Placebo</td>
<td>Expect Active</td>
<td>Expect Placebo</td>
</tr>
<tr>
<td>Compliance</td>
<td>72.9 (.30)</td>
<td>69.2 (.44)</td>
<td>67.1 (.40)</td>
<td>69.2 (.43)</td>
</tr>
<tr>
<td>Counselling</td>
<td>4.2 (.31)</td>
<td>4.4 (.46)</td>
<td>3.8 (.41)</td>
<td>3.2 (.44)</td>
</tr>
<tr>
<td>SOCRATES</td>
<td>25.9 (.60)</td>
<td>25.9 (.91)</td>
<td>25.6 (.87)</td>
<td>27.3 (1.0)</td>
</tr>
</tbody>
</table>
dependence with relevant baseline characteristics controlled for. There was, however, a marginally non-significant trend for participants who believed they received active treatment to attend more counselling sessions than those who believed they received placebo, \( F(1,115)=3.64, p=.06, \) all others \( F<1. \)

### 3.4 Discussion

As previously reported (Morley et al., 2006) there were no differences between those receiving the active treatment and those receiving the placebo. The new finding was that participants who believed they were given active treatment reported drinking less, being less alcohol dependent, and craving less alcohol than those who believed they were given a placebo. These effects were substantial, especially for total alcohol consumption. Here, participants who believed they received active treatment reported consuming half as many alcoholic drinks as participants who believed they received placebo. These effects were also not limited to positive outcomes. Participants who believed they received the active treatment also reported more adverse side effects than participants who believed they received a placebo. The lack of interaction between actual and perceived treatment for these outcomes suggests that the effect of perceived treatment was independent of the actual treatment participants were receiving. In fact, the only significant interaction between actual and perceived treatment was for total number of days abstinent.

These findings are consistent with previous studies that have shown that perceived treatment is related to actual treatment responses (Bausell et al., 2005; Dar et al., 2005; Lewis et al., 1975; McRae et al., 2004; Thomas et al., 2008). The current reanalysis was, however, unique in that blinding was maintained and there was no
effect of the active treatment compared with placebo. As a result, actual and perceived
treatment were likely to have been independent of each other. Furthermore, the effects
found here were after controlling for motivation to change, which suggests that
motivation cannot account for the relationship between perceived treatment and actual
treatment responses. This point is strengthened by the lack of significant differences
observed between those who believed they received active treatment and those who
believed they received placebo for compliance, number of counselling sessions
attended, and posttreatment motivation.

There was, however, a fundamental limitation to this reanalysis. Because
participants’ beliefs about their treatment were assessed retrospectively, it is
impossible to determine the direction of causality between perceived treatment and
actual treatment responses. As discussed in Chapter 2, participants who believe they
are receiving active treatment are much more likely to expect improvement than those
who believe they are receiving a placebo, and this may lead to greater improvement
and adverse side effects in the former via the placebo effect. On the other hand,
participants who noticed improvement as a result of spontaneous remission or random
fluctuations in their symptoms may have simply concluded that they were taking
active treatment. The same applies for adverse side effects. Importantly, this is a
limitation that applies to all retrospective reanalyses of perceived treatment in double-
blind RCTs. A second limitation is that all outcomes assessed here were self reported
and therefore subject to participant bias. Particularly concerning is the possibility for
demand characteristics, which could have contributed to the observed differences in
outcomes between those who believed they received active treatment and those who
believed they received a placebo. This is also a limitation that applies to all previous
reanalyses, with the exception of Thomas et al. (2008), who verified the higher rates of smoking cessation in participants who believed they received active treatment by assessing carbon monoxide and cotinine.

Overall, this reanalysis adds to the literature by confirming the strong relationship between perceived allocation and actual treatment responses, in this case, for treatment for alcohol dependence and adverse side effects. The success of blinding and the active treatment’s lack of efficacy suggest that these findings were not contaminated by a strong relationship between active and perceived treatment. The inability to determine the causal direction of the relationship between perceived allocation and treatment outcome highlights an important limitation to these types of studies as does the potential for participant bias, particularly demand characteristics. In the next chapter, I address these limitations by using feedback to experimentally manipulate participants’ perceived treatment and incorporating objective outcome variables in two dummy (placebo only) double-blind RCTs for cognitive performance. In later chapters, I also explore the relationship between expectancy and adverse side effects.
Chapter 4 – Perceived Treatment, False Feedback, and Cognitive Performance: Two Dummy Double-Blind RCTs (Study 2)

4.1 Introduction

The reanalysis conducted in the previous chapter demonstrated a strong relationship between perceived treatment and actual treatment responses in a double-blind RCT for alcohol dependence. This is consistent with other similar reanalyses (Bausell et al., 2005; Dar et al., 2005; Lewis et al., 1975; McRae et al., 2004; Thomas et al., 2008). The two advantages of my reanalysis were that blinding was maintained, meaning that there was likely to be a smaller relationship between actual and perceived treatment, and the ability to control for motivation to change. However, it also highlighted two important limitations that apply to these types of studies. They are: 1) the inability to determine the direction of causation between perceived treatment and actual treatment responses and 2) the predominant use of self reported outcomes that allow for participant bias, particularly demand characteristics.

While the problem of self reported outcomes can be overcome by the inclusion of objective outcome measures, the inability to determine the causal nature of the relationship between perceived treatment and actual treatment responses reflects a more significant problem. This concerns whether participants improve because they believe they have been given active treatment or whether they believe they have been given active treatment because they improved. The reanalyses conducted in Chapter 3 and elsewhere cannot differentiate between these two possibilities because they are retrospective. They make no attempt to manipulate participants’ perceived treatment
but simply measure the association between these beliefs and participants’ treatment responses. Although there is also evidence that perceived treatment appears to be influenced by improvement (see Shapiro & Shapiro, 1997b for a review), this is also retrospective and is equally insufficient. The only difference between this and the reanalyses mentioned above is the choice to use treatment response as the independent variable and perceived treatment as the dependent variable. In essence, they are assessing the same relationship but from a different point of view.

By far the biggest hindrance to determining the direction of causation between perceived treatment and actual treatment responses is the lack of experimental studies in this area. The only experimental studies that I am aware of which consider perceived treatment have compared instructions given as in double-blind RCTs with those given in standard clinical practice (Hughes et al., 1989; Kirsch & Rosadino, 1993; Kirsch & Weixel, 1988; Nash et al., 2002; Pollo et al., 2001). To my knowledge, there have been no experimental studies that have assessed whether differences in perceived treatment lead to differences in actual treatment responses within double-blind RCTs.

With this in mind, the following two experiments were aimed at developing an experimental model in which the relationship between perceived treatment and actual treatment responses could be examined more closely. This involved dummy double-blind RCTs in which participants received placebo treatment under the guise of a double-blind RCT for cognitive performance. The first advantage of this was that any potential confound between active treatment and perceived treatment was eliminated because all participants received placebo. The second advantage was that the feedback
participants received about their cognitive performance could be manipulated. This meant that I could directly test whether observable improvement, albeit bogus, influenced participants’ perceived treatment. The third advantage was that I could assess whether perceived treatment predicted cognitive performance beyond any effect of the feedback manipulation itself. In the second experiment, this was taken further by comparing the influence of feedback in participants undergoing the dummy double-blind trial, that is when it was likely to induce beliefs about treatment allocation, with another group of participants who were told they were controls in a trial for cognitive performance and received no treatment, where no such beliefs could be expected. This meant that I could ensure that any effect of perceived treatment was due to participants’ beliefs, rather than the feedback manipulation itself.

4.2 Experiment 1

Experiment 1 was a dummy double-blind RCT of caffeine for cognitive performance. Participants were asked at the end of the trial whether they believed they had been given caffeine or a placebo and asked to rate their confidence regarding this belief. There were three main points of interest. The first was whether participants who received false feedback indicating improvement as a result of their treatment would be more likely to believe they had consumed caffeine than those who received false feedback indicating no change as a result of their treatment. The second was whether perceived treatment predicted actual cognitive performance, independently of any effect of the false feedback. The third was whether confidence ratings regarding perceived treatment accounted for more variation in cognitive performance than did forced choice perceived allocation, if perceived treatment did predict cognitive performance.
4.3 Methods

4.3.1 Participants

Thirty-one (18 female) first year psychology students from the University of Sydney participated in order to gain course credit. The participants had a mean age of 18.7 (SD=.98) and drank an average of less than one cup of caffeinated coffee per day (mean=0.85, SD=.80).

4.3.2 Design

Experiment 1 involved a single factor, 2 level design. Participants were told that they were taking part in a double-blind RCT of caffeine for cognitive performance (see Appendix 3a), but were all given a benign placebo (lactose). The experimental manipulation was the false feedback participants were given about their performance before and after receiving the placebo. As shown in Figure 4.1, participants completed a cognitive task three times. The first represented their baseline performance, the second represented their performance after treatment but before feedback (pre-feedback), and the third reflected their performance after both treatment and the feedback manipulation (post-feedback). Half of the participants were given feedback indicating that their cognitive performance had improved from baseline after receiving treatment on the two subsequent tests (positive feedback). The other half of the participants received feedback indicating no change from baseline as
Figure 4.1. Design for Experiment 1. Participants in the positive feedback group received false feedback indicating that they improved by 20% from baseline after the treatment was delivered on the two subsequent tests. Participants in the no change feedback group received false feedback indicating that there was no change in their cognitive performance after the treatment was delivered.

a result of their treatment (no change feedback). In both cases the feedback was independent of the participants’ actual performance. After the final cognitive task, participants were asked about their perceived treatment. The study was single-blind and participants were randomly allocated to either positive or no change feedback.

4.3.3 Materials

Placebo Pills: The placebo pills were made from lactose and were white coated. They were prepared by the Faculty of Pharmacy, University of Sydney.

Cognitive Task: The cognitive task used rapid visual information processing (RVIP) to assess sustained attention. This was based on the RVIP task used by Yeomans, Ripley, Davies, Rusted, and Rogers (2002) who found that caffeine improved accuracy and decreased reaction times. Numbers ranging from 1-8 were presented in the middle of a computer screen at a rate of 120/min. The participants
were required to press the space bar as quickly as possible as soon as they identified either three consecutive even or three consecutive odd numbers. Participants had 1.5s to make a correct response with all responses outside this time considered false alarms. Each version of the task lasted 5min and the sequence of numbers was semi-random such that there were 8 targets every minute. The programme recorded the number of correct responses (hits), the number of incorrect responses (false alarms; FA), and the reaction time for a correct response. A non-parametric measure of accuracy was calculated, namely \( \frac{p(\text{Hit}) \cdot p(\text{FA})}{1 - p(\text{FA})} \) (McNicol, 1972), so that a single score considered both hits and false alarms. To avoid the influence of outliers, the median reaction time for a hit was taken as the measure of reaction time.

**Feedback:** The feedback manipulation was incorporated into the cognitive task. At the end of each test participants received feedback regarding their performance. This was in the form of a bogus percentage accuracy score such that all baseline scores varied randomly from 40-60%. Then, those receiving positive feedback received scores indicating a 20±2% increase from baseline on the two posttests while participants receiving neutral feedback received scores indicating no change from baseline, that is a 0±2% change.

**Beliefs Questionnaire (Appendix 3b):** This was a study-specific questionnaire containing a forced choice question that asked participants whether they believed they had received caffeine or placebo. This question read ‘Do you think that you were given the caffeine pill or the placebo pill?’. The next question asked participants to rate how confident they were about their perceived medication using an 11-point (0 to 10) Likert-type scale. This read ‘On a scale from 0-10, how certain are you that this is
the treatment you received?’. The scale was anchored by ‘Not at all’, a confidence rating of 0, and ‘Extremely certain’, a confidence rating of 10. These confidence ratings served two purposes. First, they allowed for the assessment of differences in the strength of confidence about perceived treatment between those who believed they had received caffeine and those who believed they had received placebo. Second, this scale was used as a predictor of cognitive performance by converting the 11-point scale to a 21-point scale (-10 to 10) that reflected perceived treatment and confidence simultaneously. This was done by allocating those who believed they received the placebo a negative score for this question while those who believe they received active treatment maintained positive scores. As such, the original 11-point scale can be thought of as the strength of perceived treatment, while the 21-point scale can be thought of as perceived treatment ratings.

4.3.4 Procedure

Participants attended a single one hour session. The information sheet (see Appendix 3a) explained that the study was a double-blind RCT of caffeine aimed at confirming that caffeine increases cognitive performance. After informed consent was gained, the experimenter described the RVIP task to the participants and gave them an opportunity to practise on two 30s trial versions with no feedback given. In order to make the feedback manipulation meaningful, participants were told that university students of the same age typically scored between 40-60% on the cognitive task. After this, the participants completed an initial 5min version of the RVIP task which constituted their baseline accuracy. At the end of this task all participants received feedback indicating accuracy between 40-60%, regardless of their actual score.
Participants were then given the placebo pill under the guise of a double-blind RCT of caffeine as a cognitive enhancer. Specifically, they were given a placebo pill packaged in a numbered envelop and told that the envelope contained either caffeine or placebo, but that they would not know which one they had been allocated to receive. They were also told that the experimenter delivering the envelope was unaware of whether it contained caffeine or the placebo. Participants were then given a 10min rest, which they were told was to allow the treatment to have its effect. During this time they were allowed to read general interest magazines. After 10min, the participants completed a second version of the RVIP task which constituted their pre-feedback performance. Upon finishing this task, half of the participants received positive feedback indicating that their cognitive performance had improved from baseline, while the other half received feedback indicating no change from baseline. Following another 10min rest period, the participants completed a third and final version of the RVIP task, which constituted their post-feedback performance. Feedback for this test mimicked that of the previous test. That is, those who received positive feedback after the pre-feedback test also received positive feedback on the third test that indicated sustained improvement from baseline. Participants who had received feedback indicating no change from baseline on the pre-feedback test, again received feedback indicating no change since baseline. Finally, participants completed the beliefs questionnaire which contained questions regarding their perceived treatment allocation. Participants were tested either alone or in pairs. These procedures were approved by the University of Sydney’s Human Research Ethics Committee.
4.3.5 Statistical analysis

A Chi-square test of independence assessed whether positive feedback led participants to believe they had been given caffeine more frequently than feedback indicating no change did. The strength of confidence regarding perceived treatment was compared using an independent samples $t$-test. Given that the feedback manipulation could itself affect cognitive performance (see Kluger & DeNisi, 1996 for a review) it was desirable to test the impact of perceived treatment on cognitive performance while controlling for feedback. However, it was not possible to include feedback and perceived treatment in a simultaneous regression in the current experiment because they were too strongly correlated and this led to problems with multicollinearity (see Cohen, Cohen, West, & Aiken, 2003 for a discussion of this problem). As a result, simple linear regression was used to examine the impact of feedback and perceived treatment on cognitive performance separately. This involved testing accuracy and reaction times across feedback and forced choice perceived treatment at 1) baseline, 2) pre-feedback while controlling for baseline, and 3) post-feedback while controlling for pre-feedback scores as well as perceived treatment ratings on accuracy and reaction times on the post-feedback test controlling for pre-feedback performance. Because they occurred before feedback the first two tests, baseline and pre-feedback, were simply to ensure that there were no initial differences in cognitive performance across groups. The important test in terms of assessing the possible impact of perceived treatment was the post-feedback test. All statistical analyses were conducted using SPSS software (version 15; SPSS Inc, Chicago, Ill) and results were considered significant when $p<.05$. 
4.4 Results and discussion

4.4.1 Feedback on perceived treatment

Figure 4.2 displays forced choice perceived treatment by feedback. Seventy-three percent of participants who received feedback indicating improvement as a result of their treatment believed they had been given caffeine compared with only 6.3% of those given feedback indicating no change, $\chi^2=14.7$, df=1, n=31, $p<.01$. This indicates that the positive feedback was generally successful at inducing participants to believe they had been allocated to receive caffeine as was the no change feedback at inducing participants to believe they had been allocated to receive the placebo.

Using the original 11-point scale, there were no differences in strength of beliefs between participants who believed they had been given caffeine (M=5.1, SD=2.4) and those who believed they had been given the placebo (M=5.9, SD=2.6), $t(29)=.88$, $p=.39$. 

![Feedback on Perceived Treatment (FC)](image_url)
4.4.2 Feedback on cognitive performance

Mean accuracy and reaction times by feedback for the three tests are presented in Figure 4.3. There were no significant differences between those receiving positive feedback and those receiving no change feedback on accuracy or reaction time on the baseline and pre-feedback tests, highest \( t(28)=1.11, p=.28 \). Importantly, there were also no differences in accuracy or reaction time between those receiving positive feedback and those receiving no change feedback on the post-feedback test controlling for performance on the previous test, highest \( t(28)=1.02, p=.35 \). This lack of a significant difference on the post-feedback test indicates that the feedback manipulation in and of itself did not predict cognitive performance. That is, there was no evidence to suggest that receiving positive feedback, in and of itself, led to better cognitive performance than receiving negative feedback did.

**Figure 4.3.** Experiment 1. Mean (±SE) accuracy (A) and reaction time (B) for those who received positive feedback and those who received no change feedback. No differences were significant at any stage.
4.4.3 Perceived treatment on cognitive performance

As shown in Figure 4.4A, for forced choice perceived treatment there were no differences in accuracy on the baseline and pre-feedback tests between those who believed they received caffeine and those who believed they received the placebo, highest \( t(28)=1.35, p=.19 \). On the post-feedback test, however, participants who believed they received caffeine scored 12.3% higher than those who believed they received placebo controlling for performance on the previous test, \( t(28)=2.26, p=.03 \).

This suggests that participants who believed they received caffeine performed better than participants who believed they received the placebo on the post-feedback test.

When confidence ratings were analysed, a 1-point increase in perceived treatment rating corresponded to a 1.2% increase in accuracy on the post-feedback test, \( t(28)=2.62, p=.01 \), which further supports the predictive power of perceived treatment on cognitive performance. The ratings uniquely accounted for 9.6% of the variance in accuracy at the second posttest, while the forced choice measure accounted for 7.4%
of the variance. Although this is only a small difference it might suggest that perceived treatment ratings are a more sensitive measure than forced choice perceived treatment. There were no differences in reaction time as a result of any measure of perceived treatment on any of the tests, see Figure 4.4B, all $F<1$, suggesting that the relationship between perceived treatment and cognitive performance was confined to accuracy.

4.5 Experiment 2

The feedback manipulation employed in Experiment 1 was successful at inducing beliefs about treatment allocation. Participants who received positive feedback after receiving treatment were much more likely to believe they had been given caffeine than those who received no change feedback. Further, participants who believed they received caffeine demonstrated higher accuracy on the final test than those who believed they received placebo. Although the feedback itself did not predict cognitive performance, Experiment 1 could not completely rule out the possibility that the false feedback contributed to some of perceived treatment’s effect on cognitive performance. This was because perceived treatment and false feedback were so strongly related.

Experiment 2 aimed to overcome this limitation by comparing the impact of false feedback when it induced beliefs about perceived treatment with false feedback when no such beliefs could be expected. This was done by adding a control group that underwent exactly the same procedures described in Experiment 1, but was told that they had been allocated to a no treatment group in an open trial for cognitive performance. If it is positive feedback in and of itself that leads to improved cognitive
performance, then participants who receive positive feedback should demonstrate better cognitive performance than those receiving no change feedback regardless of whether they receive placebo treatment disguised as a double-blind RCT or whether they are told they are controls in an open trial and receive no treatment. On the other hand, if it is the belief about being on active treatment (which the positive feedback induces) that leads to improved cognitive performance, then those who believe they have been given active treatment because they received positive feedback should demonstrate better cognitive performance than those who believe they have been given placebo because they receive no change feedback. In this case, minimal, if any, differences between positive and no change feedback would be expected in those receiving no treatment because there are no such beliefs about treatment allocation.

A second, more minor modification was the use of piracetam as the bogus active treatment rather than caffeine. Caffeine is the world’s most commonly used psychoactive substance (Fredholm, Battig, Holmén, Nehlig, & Zvartau, 1999). This means that participants are likely to have strong preconceived expectancies about caffeine’s effects. Piracetam on the other hand is a nootropic which is likely to be relatively unknown by first year psychology students and should elicit less preconceived expectancies about its effects. If so, using piracetam rather than caffeine as the bogus active treatment should make the information provided in the experiment, including the feedback manipulation, more salient because participants are less likely to enter the experiment with preconceived notions about its effect.
4.6 Methods

Except where stated otherwise, the methods used in Experiment 2 were identical to those of Experiment 1.

4.6.1 Participants

Forty-eight (33 female) first year psychology students from the University of Sydney participated in order to gain course credit. The participants had a mean age of 18.8 (SD=1.0).

4.6.2 Design

The main change in Experiment 2 was the addition of a control group that received the same feedback manipulation but received no treatment. This involved a 2x2 design with feedback and treatment as factors. The design is shown in Figure 4.5. Participants were told that they were taking part in a double-blind RCT of piracetam for cognitive performance, but, as with Experiment 1, all participants actually received a benign placebo (lactose). Also as with Experiment 1, half of the participants received bogus feedback indicating that their cognitive performance improved from baseline while the other half received bogus feedback indicating no change from baseline. The new aspect of this experiment was the treatment manipulation. Participants were randomly allocated to receive the placebo pill under double-blind instructions (double-blind group; see Appendix 3c) or they were told
that they were controls in a trial of piracetam and would not be receiving treatment (no treatment group; see Appendix 3d). This randomisation was done with a 3:2 ratio such that 60% of participants were allocated to the double-blind group and 40% of participants were allocated to the no treatment group.

4.3.3 Materials

*Placebo Pills:* The placebo pills were still made from lactose but were red coated in Experiment 2. They were also prepared by the Faculty of Pharmacy, University of Sydney.

*Beliefs Questionnaire (Appendices 3e and 3f):* was identical to Experiment 1 except that participants were asked whether they believed they had received piracetam or placebo rather than caffeine or placebo for those in the double-blind group. Those
in the control group were simply asked to note whether they had been given piracetam or no treatment.

4.3.4 Procedure

The procedure was identical to Experiment 1 except that participants were now randomly allocated to receive placebo treatment under the guise of a double-blind RCT of piracetam or to receive no treatment under the guise of being controls in a trial of piracetam. Because of this, participants were randomly allocated to the treatment condition in pairs so that when two people were tested simultaneously they were in the same treatment group.

4.3.5 Statistical analysis

There were two parts to the statistical analysis. The first sought to replicate the findings from Experiment 1 by testing the impact of the feedback manipulation on perceived treatment and whether feedback and/or perceived treatment predicted cognitive performance in only those who received the double-blind placebo treatment. The analysis conducted to achieve this was identical to that of Experiment 1.

The second part examined whether perceived treatment or the feedback manipulation affected cognitive performance. As with the analysis above, the important test in terms of assessing this was the post-feedback test, because this was after the first feedback manipulation. Therefore, ANCOVA assessed the impact of feedback and treatment on cognitive performance on the post-feedback test controlling for scores from the pre-feedback test. Any significant interactions found were investigated further by tests of simple effects using Fisher’s LSD procedure.
This analysis included all participants in the no treatment group, but only included participants in the double-blind group whose perceived allocation matched their feedback, that is, believed piracetam after receiving positive feedback or believed placebo after receiving negative feedback. This was done because the critical question was whether positive feedback led to better cognitive performance regardless of treatment type or whether the belief about being on piracetam induced by positive feedback led to better cognitive performance. By excluding the 5 (17%) participants in whom the feedback failed to induce matched beliefs about perceived allocation, the only difference between positive feedback in the double-blind group and positive feedback in the no treatment groups was that the former believed they had been given an active treatment while the latter knew they were not receiving any treatment. Similarly, the only difference between no change feedback across treatment groups was that participants in the double-blind group believed they were taking placebo and participants in the no treatment group knew were not receiving treatment. As such, this meant that I could directly test whether the effects of positive treatment on cognitive performance, if any, were constant across treatment or whether improvement only occurred when participants believed they had been given active treatment. All statistical analyses were conducted using SPSS software (version 15; SPSS Inc, Chicago, Ill) and results were considered significant when $p<.05$.

4.7 Results and discussion

The following 3 sub-sections (4.7.1-4.7.3) refer to the analysis conducted on the double-blind group only. This analysis tested whether the results found in Experiment 1 were replicable.
4.7.1 Feedback on perceived treatment

As shown in Figure 4.6, based on the forced choice question, 86% of participants in the double-blind group who received positive feedback believed they had been given piracetam, the bogus active treatment, while only 19% of those who received no change feedback believed they had been given piracetam, \( \chi^2=13.4, \text{df}=1, n=30, p<.01 \). This replicates the results of Experiment 1 and indicates that feedback, albeit false, influences perceived treatment. Using the original 11-point scale, participants who believed they had been given piracetam had a mean confidence of 5.2 (SD=2.0) while those who believed they had been given the placebo had a mean confidence of 4.9 (SD=2.6) which were not significantly different, \( t(28)=.33, p=.75 \).

**Figure 4.6.** Experiment 2. Feedback on forced choice (FC) perceived treatment. Participants who were given positive feedback about their cognitive performance were significantly more likely to believe they had been given piracetam compared with participants who were given no change feedback.

4.7.2 Feedback on cognitive performance

Mean accuracy and reaction times by feedback for only those in the double-blind group are shown in Figure 4.7. As with Experiment 1, controlling for
performance on the previous test where applicable, there were no significant
differences in accuracy or reaction times between participants who received positive
feedback and those who received no change feedback, highest \( t(27)=1.17, p=.25 \).
There were also no significant differences in accuracy, \( t(27)=1.54, p=.14 \), or reaction
times as a function of feedback, although the tendency towards faster reaction times in
those who received positive feedback compared with those who received no change
feedback was marginally non-significant, \( t(27)=1.82, p=.08 \). This also replicates the
findings of Experiment 1 and indicates that feedback manipulation in and of itself is
not a significant predictor of cognitive performance.

4.7.3 Perceived treatment on cognitive performance

Figure 4.8 displays mean accuracy and reaction time by perceived treatment.
Also as with Experiment 1, controlling for performance on the previous test as
appropriate, there were no differences in accuracy or reaction times between those
who believed they received piracetam and those who believed they received placebo
on the baseline and pre-feedback tests, highest \( t(27)=1.10, p=.32 \). Unlike Experiment
1, there was no difference in accuracy on the post-feedback test between those who believed they received piracetam and those who believed they received placebo on either the forced choice or 21-point ratings of perceived treatment, both $F<1$. This reflected a failure to replicate the finding from Experiment 1 that perceived allocation predicted accuracy on the cognitive task. However, participants who believed they received piracetam had significantly faster reaction times on the post-feedback test than those who believed they had been given the placebo, MD=34.5, $t(27)=2.18, p=.04$.

Similarly, a 1-point increase in perceived treatment ratings significantly predicted a 3ms decrease in reaction times, $t(27)=2.18, p=.04$. This suggests that perceived treatment did significantly predict one aspect of cognitive performance, albeit different to the one predicted in the previous experiment. Possible reasons for this difference are discussed further below. There were minimal differences in the proportion of variation that the forced choice question and ratings accounted for in reaction times, 9.7% and 9.3% respectively, suggesting that there was no real difference in their predictive power.

Figure 4.8. Experiment 2. Mean (±SE) accuracy (A) and reaction time (B) based on participants forced choice perceived treatment. On the post-feedback test, participants who believed they received piracetam had significantly faster reaction times than those who believed they received the placebo. No other differences were significant.
4.7.4 Perceived treatment versus feedback

Mean accuracy and reaction time by feedback for those in the double-blind group and those in the no treatment group are shown in Figure 4.9. Controlling for performance on the previous test, the two-way ANCOVA revealed that there were no main effects of either perceived treatment or feedback on the post-feedback test for accuracy and for reaction time, all $F<1$. The interaction for accuracy was also non-significant, $F(1,38)=1.07$, $p=.3$. There was, however, a significant interaction between perceived treatment and feedback for reaction time, $F(1,38)=8.86$, $p<.01$. This indicated that the differences between positive and no change feedback were not the same for those in the double-blind group as they were for those in the no treatment group. Test of simple effects revealed that, for the no treatment group, positive feedback actually led to significantly slower reaction times than did negative feedback, $F(1,38)=4.36$, $p=.04$. For the double-blind group, however, positive feedback led to significantly faster reaction times than those who received negative feedback, $F(1,38)=4.57$, $p=.04$.

The crucial difference between feedback in the double-blind group and the no treatment group was that it induced beliefs about perceived treatment in the former. That is, participants who were told they were taking part in a double-blind RCT and received positive feedback believed they had been given piracetam, the bogus active treatment, and those that received no change feedback believed they had been given a placebo. No such beliefs could be expected in the no treatment group because these participants knew that they were not receiving treatment. This suggests that the faster reaction times for those who received positive feedback and believed they were on
active treatment compared with those who received negative feedback and believed they were on placebo resulted from differences in beliefs, not feedback. Without beliefs about treatment allocation, positive feedback actually led to slower reaction times than negative feedback.

\[\text{Figure 4.9 Experiment 2. Mean (+SE) accuracy (A) and reaction time (B) on the post-feedback test by feedback for those who were in the double-blind group and those in the no treatment group. Test of simple effects revealed that positive feedback led to significantly faster reaction times in the double-blind group but actually increased reaction times in the no treatment group.}\]
4.8 General Discussion

This study used dummy (placebo only) double-blind RCTs for cognitive performance to explore the relationships between feedback, perceived treatment, and actual treatment responses. In both experiments, participants’ perceived treatment was influenced by the feedback they received. Participants who received feedback indicating they had improved as a result of their treatment generally believed they had been given active treatment while those given feedback indicating no improvement generally believed they had been given a placebo. If the positive feedback used here is a sufficient analogy for improvement that participants may experience during real double-blind RCTs, then these findings clearly demonstrate that observable improvement leads participants to believe they are on active treatment much more frequently than when they observe a lack of improvement. While this is consistent with correlational evidence that improvement is related to perceived treatment (e.g. Margraf et al., 1991; Morin et al., 1995), this is the first study to show experimentally that observable improvement does influence perceived treatment.

Participants who believed they received active treatment demonstrated better cognitive performance than those who believed they received the placebo. This is consistent with my reanalysis of a double-blind RCT for alcohol dependence in Chapter 2 and those previously conducted in other areas (Bausell et al., 2005; Dar et al., 2005; Lewis et al., 1975; McRae et al., 2004; Thomas et al., 2008). Feedback did not significantly predict cognitive performance in either experiment in the participants who received placebo treatment under the guise of a double-blind RCT. This seemed to suggest that perceived treatment was the most important determinant of actual treatment responses. That is, it seemed to be the belief that one was on active
treatment that led to better cognitive performance rather than simply receiving positive feedback.

The addition of a control group that received the same feedback manipulation but that received no feedback in Experiment 2 provided the most important finding. That is, while positive feedback led to better cognitive performance than negative feedback after placebo treatment administered under the guise of a double-blind RCT, the opposite was true when no treatment was given. This suggests that the differences in cognitive performance between those who believed they received active treatment and those who believed they received the placebo cannot be explained by the different feedback they received. Instead, it seems likely that it was the participants’ beliefs about their treatment allocation that influenced their cognitive performance. This supports a placebo based interpretation of the relationship between perceived treatment and actual treatment responses. That is, participants who believed they were taking active treatment expected to perform better than those who believed they were given the placebo and these expectations for greater improvement led to better cognitive performance. Importantly, the observed improvement was for an objective outcome which greatly reduces the possibility that these differences arose from participant bias and further supports a placebo effect based interpretation.

There seemed to be minimal differences between the use of forced choice and ratings of perceived treatment. This contradicted suggestions that confidence ratings might capture more variation in beliefs about treatment allocation than the commonly used forced choice questions (e.g. Margraf et al., 1991; Sharpe et al., 2003). However, the marginal increase in variability accounted for by the ratings in Experiment 1 may
imply that it is worth including these when assessing perceived treatment given how easily they can be incorporated.

There are a number of important implications of these findings. The finding that observable changes influenced perceived treatment reinforces Sharpe et al.’s (2003) concern that the probability of blinding being broken is likely to increase with the magnitude of the treatment’s efficacy. This presents a significant problem for double-blind RCTs because it produces an unsatisfactory situation whereby these trials cannot validly assess efficacious treatments. This is because the more efficacious the treatment is the more likely it will produce improvement that enables participants to determine whether they have been allocated to active treatment or placebo. Another important implication concerns the interplay between observable changes, perceived treatment, and actual treatment responses. The possibilities that either observing improvement causes perceived treatment or that perceived treatment causes improvement need not be mutually exclusive. In the current study false feedback suggesting either improvement or no change strongly influenced perceived treatment and perceived treatment, in turn, appeared to influence actual treatment responses. This suggests that observing some improvement may trigger a belief about being on active treatment which may cause more improvement via the placebo effect.

There are also some potential limitations to the current study. The fact that perceived treatment significantly predicted accuracy in Experiment 1 but predicted reaction times in Experiment 2 might raise questions regarding the true replicability of these findings. Importantly, this difference between Experiments 1 and 2 cannot be explained by processes such as speed-accuracy trade-off. In Experiment 1, when
perceived treatment predicted accuracy there were no significant differences in reaction times. Similarly, in Experiment 2, when perceived treatment predicted reaction times there were no differences in accuracy. One possible explanation for the difference is that changing the bogus active treatment from caffeine (Experiment 1) to piracetam (Experiment 2) created differences in how the participants’ expected their cognitive performance to be affected. When caffeine was the bogus active treatment participants may have expected their accuracy to improve, whereas when piracetam was the bogus active treatment participants may have expected their reaction times. However, because I did not assess participants expectancies regarding the efficacy of their treatment I was unable to test this possibility.

As with many other studies (e.g. Basoglu, Marks, Livanou, & Swinson, 1997; Bausell et al., 2005; Dar et al., 2005; Rabkin et al., 1986), perceived treatment was only assessed after the outcome of interest was measured, in this case cognitive performance on the post-feedback test. Although this was done intentionally in an attempt to ensure that participants did not question the purpose of the study, it remains possible that some participants may have noticed improvement on the post-feedback test and that this led them to believe they had been allocated to receive the active treatment. However, given the success at which feedback induced beliefs about treatment allocation, this possibility seems to rest upon unsuccessful randomisation or highly spurious increases in cognitive performance which coincided with receiving positive feedback.

Overall, this study provides firm evidence that participants in double-blind RCTs use available cues, in this case feedback, in order to determine whether they
have been allocated to receive active treatment or a placebo. Although this has been suggested by correlational evidence from previous studies (see Shapiro & Shapiro, 1997b for a review) this is the first study to show this effect by experimentally manipulating observable changes via false feedback. Further, there was a clear relationship between perceived treatment and cognitive performance. Participants who believed they received active, albeit bogus, treatment performed better than participants who believed they received the placebo and these differences could not be explained by the feedback manipulation itself or self report bias. This suggests that perceived treatment can influence outcomes in double-blind RCTs via the placebo effect. Finally, the use of dummy (placebo only) double-blind RCTs also ruled out the potential for confounding relationships between active treatment and perceived treatment and may serve as a very useful model for exploring these issues further.

Attention now turns to the possible contribution of the placebo effect to treatment side effects.
Chapter 5: Review of Placebo-Induced Side Effects

The earlier finding that participants taking part in a double-blind RCT for alcohol dependence who believed they were given active treatment reported more side effects than participants who believed they were given a placebo (see Chapter 3) suggests that the placebo effect may contribute to treatment side effects. While this possibility has been raised by a number of other researchers (e.g. Barsky et al., 2002; Myers, Cairns, & Singer, 1987; Roscoe et al., 2006; Shapiro, Chassan, Morris, & Frick, 1974), it has received surprisingly little research attention, with the majority of placebo studies focusing on beneficial outcomes.

In addition to being inherently unpleasant, adverse side effects often lead to poorer compliance with and in some cases discontinuation of treatment, which, in turn, leads to worse treatment outcomes (e.g. chemotherapy: Demissie, Silliman, & Lash, 2001; antidepressants: Schatzberg, 2007). Determining if and how the placebo effect contributes to these adverse effects may, therefore, be useful in guiding interventions to reduce the burden of side effects and improve treatment outcomes.

The limited evidence for placebo-induced side effects to date can be categorised into: 1) experimental studies assessing negative placebo effects, 2) adverse side effects in placebo groups from double-blind RCTs, 3) studies assessing the relationship between pretreatment expectancies and post-chemotherapy nausea, and 4) experimental studies assessing whether warning patients about side effects actually leads to more adverse side effects. This chapter reviews the aetiology and evidence for placebo-induced side effects.
5.1 Possible Causes of Placebo-Induced Side Effects

As discussed in Chapter 1, placebo-induced side effects are any responses to a treatment other than those for which the treatment has been administered and that are not attributable to the inherent properties of the treatment itself. Placebo-induced side effects are, therefore, best considered as a subset of the placebo effect, with the same underlying mechanisms. In this way, placebo-induced side effects are learned responses that could result from information and/or classical conditioning. Note that this is not to say that placebo-induced side effects and all other placebo effects are produced by the same biological processes, as this is not even the case for positive placebo effects. For example, Amanzio and Benedetti (1999) found that placebo analgesia induced by verbal information could be reversed by naloxone, while placebo analgesia induced by conditioning with keterolac was unaffected by naloxone, suggesting that the former but not the latter was mediated by endogenous opioids. Instead, it simply means that placebo-induced side effects could occur as a result of the information patients receive about their treatment’s side effects as well as any previous experience they have had with that treatment or even with similar treatments.

Patients in standard clinical practice and participants in double-blind RCTs receive information about their treatment that usually includes a warning about possible side effects. This is important so that the patient or participant can make an informed decision about whether he or she wishes to receive the treatment, known as the process of informed consent. There is, however, the possibility that this information might cause patients and participants to expect and therefore experience the side effects they have been warned about (Barsky et al., 2002; Myers et al., 1987;
Shapiro et al., 1974). Thus, informing patients or participants about potential side effects may cause placebo-induced side effects.

Patients are also likely to draw on information from sources other than their health professionals, including their family and friends, the media, and the internet. This may be particularly so for more prevalent illnesses and their treatments. For example, there are innumerable internet sites that provide information about chemotherapy and its side effects and chemotherapy-induced alopecia and nausea are often portrayed on television and in movies. This information may also lead to placebo-induced side effects if it elicits sufficiently strong expectancies for these effects.

Patients will often have had prior experience with their treatment or other similar treatments and this provides opportunities for classical conditioning that may also lead to placebo-induced side effects. The clearest example of classical conditioning producing a side effect-like response is anticipatory nausea. Studies in this area have shown that the severity of post-chemotherapy nausea predicts the severity of anticipatory nausea (Andrykowski & Redd, 1987; Stockhorst, Klosterhalfen, Klosterhalfen, Winkelmann, & Steingrueber, 1993) and that anticipatory nausea is subject to processes of classical conditioning, such as overshadowing (Stockhorst et al., 1998). This means that evaluating placebo-induced side effects requires consideration of both the information the patient or participant has received about their treatment and any prior experience they have had with the treatment or with similar treatments.
5.2 Evidence for Placebo-induced Side Effects

5.2.1 Experimental studies on negative placebo effects

As with positive placebo effects, most research on negative placebo effects has been on pain. These studies have shown that placebo administration with the suggestion of hyperalgesia can cause increased pain sensitivity compared with natural history (e.g. Benedetti, Amanzio, Casadio, Oliaro, & Maggi, 1997; Benedetti, Amanzio, Vighetti, & Asteggiano, 2006; Colloca, Sigaudo, & Benedetti, 2008; Kong et al., 2008), with one exception (Johansen, Brox, & Flaten, 2003). For example, Colloca et al. (2008) told healthy volunteers that a sham electrode would increase the intensity of a painful stimulus and that activation of this electrode was signalled by a green screen, whereas deactivation was signalled by a red screen. The participants consistently rated their pain as higher when they saw the green screen compared with the red screen, even though the actual pain intensity was kept constant. The lack of evidence for placebo hyperalgesia in Johansen et al.’s (2003) probably resulted from the high intensity of the painful stimulation they used. In their study, average pain ratings for the natural history were over nine out of ten, which likely led to a ceiling effect whereby any additional pain in the negative placebo group could not be detected.

There is also some preliminary evidence to suggest that such placebo hyperalgesia is mediated by cholecystokinin systems in the opposite direction to placebo analgesia (Benedetti et al., 1997; Benedetti et al., 2006). Although far from conclusive, this seems to indicate that the effect has physiological underpinnings and exists beyond subjective bias. Information has also been shown to induce placebo headaches, which is perhaps an extension of placebo hyperalgesia. In these studies,
participants told that a sham electrode applied to their head will cause headaches reported headaches twice as often as those told that the electrode would not be switched on (Bayer, Baer, & Early, 1991; Schweiger & Parducci, 1981).

There is some evidence of placebo allergic reactions. In perhaps the most striking study on negative placebo effects, Ikemi and Nakagawa (1962) found that Japanese men who were allergic to lacquer trees reacted to resin from harmless trees when they were told that the resin was from a lacquer tree. The adverse reactions to the normally harmless trees were quite severe, with participants developing skin irritation and rashes that lasted for up to 11 days. Unfortunately there appears to have never been an attempt to replicate this study. More robust evidence comes from Luparello and colleagues (Luparello, Lyons, Bleecker, & McFadden, 1968; McFadden, Luparello, Lyons, & Bleecker, 1969) who have shown twice that informing asthmatics that a placebo inhalant, nebulised saline, contains an allergen leads to bronchoconstriction. Interestingly, they also found that this placebo bronchoconstriction could be reversed by administration of the same placebo inhalant, but with suggestion that it is a bronchorelaxant. The fact that the placebo effect can cause allergic reactions may not be entirely surprising given evidence of immunooconditioning studies which produce placebo-like effects (see Pacheco-López et al., 2006 for a review). It is also worth noting that in both these cases the participants had a history of allergic reactions and that this experience likely contributed to the placebo effects observed.

In other areas, Flaten et al. (1999) found that both a stimulant and a placebo pill led to higher ratings of tension in participants who were told to expect stimulation
in the form of increased metabolism and bodily activity compared with those given no information about the treatment. This suggests that the placebo effect can increase the magnitude of an adverse response to an active treatment as well as inducing an adverse response to an otherwise inactive treatment. Similarly, studies using the ‘open versus hidden’ design (see Chapter 2) have shown that people with Parkinson’s disease show poorer motor performance when they are aware that stimulation of their subthalamic nucleus has ceased compared with when they are unaware of this (Benedetti, Maggi et al., 2003; Mercado et al., 2006; Pollo et al., 2002). In this case, the information that the stimulation has ceased increases the decline in performance associated with the lack of stimulation. Again, this implies that the placebo effect can worsen adverse treatment outcomes and points towards negative placebo effects.

Overall, these studies are important for showing that suggestion of adverse outcomes can lead to negative placebo effects, whether as a result of this information alone or previous experience. However, to a certain extent, they provide only indirect evidence for placebo-induced side effects. This is because they involve suggestion about a sole action of the treatment being delivered, whether active or placebo. In the pain studies for example, a placebo is administered and hyperalgesia is suggested. This is somewhat different to placebo-induced side effects, which are placebo effects in response to the treatment, other than those for which it has been given. In this sense, these studies do not necessarily indicate that information or experience with treatment side effects can cause placebo-induced side effects. Instead, these types of negative placebo effects might be considered as primary placebo effects, whereas placebo-induced side effects might be conceived of as secondary placebo effects. As a
result, they provide important evidence that negative placebo effects do occur, but they do not directly test for placebo-induced side effects.

5.2.2 Side effects in placebo groups in double-blind RCTs

One source of evidence often cited for placebo-induced side effects are placebo groups in double-blind RCTs, as it quite common for these participants to report side effects. In the double-blind RCT reanalysed in Chapter 3, 48% of participants in the placebo group reported at least one side effect. The most common of these side effects was headaches, which was reported by 20% of participants receiving placebo (Morley et al., 2006). Interestingly, the occurrence of headaches was actually significantly higher in the placebo group than it was in the group receiving acamprosate (4%), suggesting possible interaction between drug effects and placebo-induced side effects. In a double-blind RCT of gabapentin for panic disorder, Pande et al. (2000) also found a high rate of side effects reported by the 51 participants receiving placebo. In all, 24% reported headaches, 18% reported somnolence, 16% reported nausea, and 14% reported asthenia and dyspepsia. These rates were very similar to the side effects reported by participants receiving gabapentin. Remarkably, two (4%) of the participants on placebo actually withdrew from the study because of these adverse effects. In another trial, Preston, Materson, Reda, and Williams (2000) found slightly lower rates of side effects in 187 participants receiving placebo in a double-blind RCT of six anti-hypertensive medications. In all, 12% of the placebo group in their study reported at least one side effect and the most common of these was again headaches. Even more remarkably than Pande et al.’s (2000) finding, 7% of the participants receiving placebo withdrew from the trial because of adverse effects.
These examples of side effects in the placebo groups of double-blind RCTs appear representative according to at least two reviews. In the first, Rosenzwieg, Brohier, and Zipfel (1993) assessed side effects in 1228 healthy volunteers allocated to receive placebos in 109 double-blind RCTs. They found that 20% of these participants reported at least one side effect in response to their placebo treatment. Headaches were the most commonly reported side effect with 7% of participants reporting headaches. They also found that participants receiving repeated placebo treatments reported more side effects than those receiving a single placebo treatment (26% vs 16%, respectively). This is interesting because it parallels evidence that more invasive treatment regimens produce larger placebo effects (de Craen et al., 1999; de Craen et al., 2000). In the second review, Weihrauch and Gauler (1999) analysed double-blind RCTs for stroke, angina pectoris, diabetes, anxiety, and gastro-duodenal lesions conducted by a pharmaceutical company. Overall, they found that between 2-62% of participants receiving placebo reported at least one side effect. These side effects also tended to mimic those reported by participants receiving active treatment.

The relatively frequent rate of side effects reported by participants receiving placebo treatment in double-blind RCTs supports the possibility that the placebo effect might contribute to treatment side effects. This is particularly so given that the side effects reported in these participants seem to mimic those reported by participants receiving active treatment, which all participants have presumably been warned about. There are, however, two important limitations to this type of evidence. Firstly, none of these trials included a no treatment control group. This means that they did not adequately control for the natural history of the condition being treated. As such, it is
impossible to determine whether the side effects reported by the participants receiving placebo directly resulted from the placebo treatment or whether they would have occurred regardless of treatment and were simply misattributed to the treatment. Therefore, the frequency of side effects found in these types of studies might overestimate placebo-induced side effect.

On the other hand, given that participant blinding is often found to be unsuccessful in these types of trials (see Chapter 2), it may be the case that they underestimate the occurrence of placebo-induced side effects. As discussed in detail in Chapter 2, participants who know they are receiving placebo treatment are much less likely to expect the treatment to affect them and may, therefore, have only weak expectancies for side effects. This makes it very difficult to determine the true extent to which side effects are associated with placebo treatment in double-blind RCTs. As a result, the occurrence of side effects in the placebo groups of double-blind RCTs is insufficient to determine whether the placebo effect contributes to treatment side effects, however suggestive it may be.

### 5.2.3 Expectancies and post-chemotherapy nausea

A number of studies have assessed the relationship between chemotherapy patients’ pretreatment expectancies for nausea and their actual experience of nausea posttreatment. Generally, these studies have employed the same basic design, involving asking first time chemotherapy patients to rate their expectancies for nausea and then to report their experience of nausea following one or more infusions. A summary of these studies and their limitations is provided in Table 5.1. Nine out of the eleven studies identified found that higher pretreatment expectancies for nausea
were associated with greater posttreatment nausea, with expectancies accounting for up to 32% of the variability in post-chemotherapy nausea. This suggests that some of the nausea chemotherapy patients experience after treatment might be attributable to the placebo effect. If so, this would provide evidence for a placebo-induced side effect. However, there are a number of methodological limitations to these studies that make it difficult to determine whether expectancies have a causal impact on post-chemotherapy nausea, or whether they are simply correlated.

From the outset, studies assessing pretreatment expectancies and post-chemotherapy nausea might be considered a weaker source of evidence for placebo-induced side effects because they are correlational in nature. However, correlational studies can provide evidence regarding the causal nature of relationships if other potentially confounding factors are adequately controlled for. To this end, the method of statistical analysis used in these studies is integral in determining whether expectancies cause, or at least worsen posttreatment nausea. Unfortunately, almost all of the studies identified here are limited in this respect, as can be seen in Table 5.1.

Firstly, the majority of these studies consisted of relatively small samples given their correlational nature, with only three involving more than 100 patients (Rhodes, Watson, McDaniel, Hanson, & Johnson, 1995; Roscoe et al., 2004; Shelke et al., 2008). Secondly, five of the studies did not control for history of nausea in other settings (Andrykowski & Gregg, 1992; Cassileth et al., 1985; Olver, Taylor, & Whitford, 2005; Rhodes et al., 1995; Shelke et al., 2008), which is problematic because expectancies for nausea might simply reflect patients’ knowledge of their risk of experiencing nausea. Thirdly, in four out of the five studies that did control for
Table 5.1. Summary of studies assessing the relationship between first time chemotherapy patients' expectancies and post-chemotherapy nausea. In all studies expectancies were assessed prior to the first infusion. In each case the patients, statistical analysis, measure of nausea, significance of expectancy, and limitations are described. Where more than one statistical test was used, multivariate analysis was favoured over univariate analysis. Other predictors included (hierarchical) or entered (stepwise) into the regression models that significantly predicted nausea are italicised. If provided by the authors, the unique proportion of variability (ΔR²) expectancy accounted for in post-chemotherapy nausea is given.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Patients</th>
<th>Statistical Analysis</th>
<th>Nausea</th>
<th>Significant</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrykowski &amp; Gregg (1992)</td>
<td>65 mixed cancer patients</td>
<td>Hierarchical regression (state anxiety)</td>
<td>Average nausea severity over 2-4 infusions</td>
<td>No</td>
<td>Small sample size, No control for history of nausea</td>
</tr>
<tr>
<td>Cassileth et al. (1985)</td>
<td>56 mixed cancer patients</td>
<td>Chi-square test of independence</td>
<td>Severity of nausea experienced prior to 3rd infusion</td>
<td>No</td>
<td>Small sample size, No control for history of nausea or any other variable</td>
</tr>
<tr>
<td>Haut et al. (1991)</td>
<td>36 breast cancer patients</td>
<td>Hierarchical regression (number of treatments, emetic potential, no. of anticancer drugs, anti-emetics, motion sickness, state anxiety)</td>
<td>a) Frequency of nausea, i.e. proportion of infusions followed by any nausea b) Average nausea severity across all infusions [No. of infusions not specified]</td>
<td>Yes – R²=.32</td>
<td>Very small sample size, Over-fitted model</td>
</tr>
<tr>
<td>Jacobsen et al. (1988)</td>
<td>45 breast cancer patients</td>
<td>Stepwise regression (age, Karnofsky status, state anxiety, trait anxiety, history of nausea for food, anxiety, pregnancy, and motion sickness, chemotherapy agent(s))</td>
<td>a) Any nausea over 6 infusions b) Frequency of nausea, i.e. reported after &lt;4 infusions versus 4 or more infusions c) Mean nausea severity over the 6 infusions d) Mean duration of nausea</td>
<td>Yes</td>
<td>Very small sample size, Used stepwise regression</td>
</tr>
<tr>
<td>Montgomery &amp; Bovbjerg (2000)</td>
<td>52 breast cancer patients</td>
<td>Stepwise regression (age, education, employment, ethnic group, marital status, tumour size, no. of positive lymph nodes, stage, chemotherapy regimen, nausea after 6th infusion, frequency of nausea for infusions 1-5, no. of side effects after infusion 6, any anticipatory nausea on any infusion)</td>
<td>Any nausea following 7th infusion</td>
<td>Yes</td>
<td>Small sample size, Used stepwise regression</td>
</tr>
<tr>
<td>Olver et al. (2005)</td>
<td>87 mixed cancer patients</td>
<td>Hierarchical regression (nurses toxicity ratings of regimen and patients' activity)</td>
<td>Severity of nausea after 1st infusion</td>
<td>Yes – R²=.05</td>
<td>Small-moderate sample size, No control for history of nausea or any other variable</td>
</tr>
</tbody>
</table>
Table 5.1. Summary of studies assessing the relationship between first time chemotherapy patients’ expectancies and post-chemotherapy nausea (continued).

<table>
<thead>
<tr>
<th>Authors</th>
<th>Patients</th>
<th>Statistical Analysis</th>
<th>Nausea</th>
<th>Significant</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhodes et al.</td>
<td>329 mixed cancer patients</td>
<td>Chi-square test of independence</td>
<td>Any nausea in the 48hrs following 1st infusion</td>
<td>Yes</td>
<td>No control for history of nausea or any other variable</td>
</tr>
<tr>
<td>(1995)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roscoe et al.</td>
<td>Study 1: 36 ovarian cancer patients</td>
<td>Hierarchical regression (emetic potential, chemotherapy agent, anti-emetics, age &lt;50, motion sickness, nausea during pregnancy)</td>
<td>a) Average nausea severity in the 60hrs following 1st and 2nd infusions</td>
<td>Yes - (R^2=.18)</td>
<td>Very small sample Over-fitted model</td>
</tr>
<tr>
<td>(2000)</td>
<td>Study 2: 86 mixed cancer patients</td>
<td>Hierarchical regression (emetic potential, no. of chemotherapy agents, anti-emetics, chemotherapy agent×5, age &lt;50, motion sickness, nausea during pregnancy)</td>
<td>b) Average nausea severity in the 60hrs following 1st and 3rd infusions</td>
<td>Yes - (R^2=.09)</td>
<td>Small sample size Over-fitted model</td>
</tr>
<tr>
<td>Roscoe et al.</td>
<td>194 breast cancer patients</td>
<td>Hierarchical regression (age, nausea during pregnancy)</td>
<td>a) Average nausea in the 5 days from 1st infusion onwards</td>
<td>No (but (p=.054))</td>
<td>Limited covariates included</td>
</tr>
<tr>
<td>(2004)</td>
<td></td>
<td></td>
<td>b) Any severe nausea (6 or 7 out of 7) in the 5 days from 1st infusion onwards</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Shelke et al.</td>
<td>322 mixed cancer patients</td>
<td>Bivariate correlation</td>
<td>a) Average nausea in the 5 days from 1st infusion onwards</td>
<td>I: Yes - (R^2=.07)</td>
<td>No control for history of nausea or any other variable</td>
</tr>
<tr>
<td>(2008)(^5)</td>
<td></td>
<td></td>
<td>I: Yes - (R^2=.09) C: Yes (-R^2=.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>b) Peak nausea rating in the 5 days from 1st infusion onwards</td>
<td>I: Yes - (R^2=.09) C: Yes (-R^2=.09)</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) Haut et al. (1991) provide insufficient information to determine which of the other variables were significant, however, they do state that motion sickness and state anxiety did not significantly predict post-chemotherapy nausea.

\(^2\) Predictors significant on at least one of the nausea outcomes.

\(^3\) No information provided about whether the other predictors were significant or not.

\(^4\) Although Montgomery & Bovbjerg (2000) assessed and analysed expectancies before the 1st infusion, they conducted their regression analysis using expectancies assessed before the 7th infusion.

\(^5\) Shelke et al.’s (2008) study was an intervention study with participants in the intervention group (I, n=159) receiving additional information about their anti-emetic treatment and those in the control group (C, n=163) receiving no additional information. Expectancies were assessed before and after information was given. Those reported here are for the pre-intervention expectancies as these were the best predictors of posttreatment nausea.
history of nausea the regression models were either over-fitted (Haut, Beckwith, Laurie, & Klatt, 1991; Roscoe, Hickok, & Morrow, 2000) or were based on stepwise procedures (Jacobsen et al., 1988; Montgomery & Bovbjerg, 2000).

Over-fitting a regression model occurs when there are less than 10-20 participants for each predictor included in the analysis. This leads to overestimation of the total amount of variability the model accounts for as well as increasing the probability of finding spurious results for individual predictors (Harrell, 2001). Thus, the significant findings in these studies may be unreliable. Stepwise regression is often erroneously used to overcome this problem (Millis, 2003). This type of regression analysis identifies the best predictor of the dependent variable without controlling for the other variables, enters this variable in a first step, and then repeats this procedure controlling for any variables entered in previous steps up until the significance of the predictor to be entered exceeds some criterion for inclusion, e.g. \( p < .10 \). Any remaining predictors that do not meet this criterion are excluded from the model. This procedure is associated with a number of problems and its use in psychology has been heavily criticised (Harrell, 2001; Mernard, 2003; Millis, 2003; Thompson, 1995). One of these problems is that stepwise regression capitalises on error variance, thereby increasing spurious findings and making the results highly sample dependent and often difficult to replicate. Further, it may lead to elimination of predictors that are theoretically important and which should be controlled for, but that do not meet the criterion for inclusion. As a result, stepwise regression can only test for prediction rather than cause and is, therefore, inappropriate for testing theories (Mernard, 2003). As such, results from the studies employing stepwise regression are useful in terms of confirming the relationship between expectancies for nausea and
actual nausea, but they are insufficient to rule out other possible causes of this relationship.

Only one study on expectancy and post-chemotherapy nausea conducted to date avoids all of these limitations. In this study, Roscoe and colleagues (2004) asked 194 first time chemotherapy patients with breast cancer to rate their pretreatment expectancies in a number of different ways after speaking with their physician. The expectancy assessment involved asking participants to rate 1) retrospectively how likely they believed they were to experience nausea before speaking to their physician, 2) how certain they were that they would experience nausea after speaking with their physician, and 3) the highest level of nausea they expected after speaking with their physician. The patients then received their first chemotherapy infusion and were given diaries to report the severity of nausea in the morning, afternoon and evening for 5 days from the initial infusion. The diaries were used to calculate average nausea across the 5 days and whether the patient experienced severe nausea, classified as a nausea rating of 6 or 7 out of 7 at any time point. Hierarchical regression controlling for age and nausea during pregnancy found that patients’ retrospective expectancies (reported in Table 5.1) significantly predicted the occurrence of severe nausea. A similar analysis, however, revealed that none of the expectancy measures were significantly associated with average nausea over the 5 days.

The two studies which had sufficiently large sample sizes, but failed to control for any other potentially confounding variables, both found significant relationships between expectancies and post-chemotherapy nausea (Rhodes et al., 1995; Shelke et
al., 2008). In the three with small sample sizes, but that were not limited by their method of statistical analysis (Andrykowski & Gregg, 1992; Cassileth et al., 1985; Olver et al., 2005), only one found a significant relationship between expectancies and post-chemotherapy nausea (Olver et al., 2005). The remaining four that employed statistical analyses that may have overestimated the importance of expectancy all found significant relationships between expectancy and posttreatment nausea (Haut et al., 1991; Jacobsen et al., 1988; Montgomery & Bovbjerg, 2000; Roscoe et al., 2000), perhaps indicative of the statistical analysis employed.

These findings are clearly inconclusive. In the most methodologically rigorous study (Roscoe et al., 2004), a relationship between expectancies and post-chemotherapy nausea was only evident for one of three methods of assessing expectancy and only for the occurrence of severe nausea. Further, although age and nausea during pregnancy were included as covariates in their analysis, susceptibility to motion sickness was not. The authors based this decision on the lack of correlation between motion sickness and posttreatment nausea. However, motion sickness remains an important variable to control for, regardless of whether or not its correlation with posttreatment nausea is significant, as it has often been found to predict nausea in other studies (e.g. Jacobsen et al., 1988; Morrow, 1985; Roscoe et al., 2000) and it may account for some of the variance in expectancies. The other, less rigorous studies seemed to provide relatively consistent evidence for an association between expectancies and post-chemotherapy nausea, but due to their limitations, they cannot determine whether expectancies actually contribute to or are merely correlated with this nausea. Hence, as with studies on the negative placebo effects and side effects in the placebo groups of double-blind RCTs, the results of the studies that have
assessed the relationship between expectancies and post-chemotherapy nausea can only be considered suggestive of placebo-induced side effects.

It is worth noting that Sohl, Schmur, and Montgomery (2009) recently conducted a meta-analysis assessing the relationship between expectancy and treatment-related side effects in cancer patients that included some studies measuring post-chemotherapy nausea. Their analysis indicated a moderate positive relationship between expectancy and adverse symptoms in these patients. However, rather strangely, they only included effect sizes for bivariate associations between expectancy and adverse symptoms, meaning that there was no control for other potentially confounding variables, such as a history of nausea in other settings. As a result, their meta-analysis can also only be considered suggestive of placebo induced side effects.

5.2.4 Experimental studies on information and side effects

Eight studies have investigated placebo-induced side effects by manipulating whether participants are warned about possible side effects or not. The experimental nature of these studies means that they are best designed to determine whether the placebo effect contributes to treatment side effects, especially because they control for natural history. These studies have produced conflicting results, with half finding that warning participants about side effects does lead to more side effects being reported (Gibbs, Waters, & George, 1989b; Mondaini et al., 2007; Myers & Calvert, 1984; Myers et al., 1987) and the other half finding no effect of the warning (Gibbs, Waters, & George, 1989a; Howland, Baker, & Poe, 1990; Morris & Kanouse, 1982; Myers & Calvert, 1973).
In terms of studies finding evidence in support of placebo-induced side effects, Myers and Calvert (1984) found that providing information only about the beneficial effects of the treatment led to less side effects being reported than when patients were warned about treatment side effects or when they received no information at all. In perhaps the most remarkable study in this area, Myers et al. (1987) compared side effects in participants in a multicentre double-blind RCT for unstable angina who were warned about the possibility of gastrointestinal discomfort with those reported by participants who were not warned about this possibility. Those who received the warning subsequently reported much higher rates of gastrointestinal discomfort and were six times more likely to withdraw from the study because of this. It is worth noting, however, that this study was only quasi-experimental because participants were not randomised to receive the warning or not. Instead, the difference in warnings arose from differences in ethical requirements between the participating centres. Two out of the three participating centres’ ethics committees required participants to be warned about possible gastrointestinal discomfort, while the others did not. Further, participants in this study received aspirin, sulfinpyrazone, or placebo and the authors did not report on whether the warning interacted with type of treatment or not. Of most interest would have been whether the warning affected participants receiving placebo treatment.

The most compelling evidence for the effect of information on side effects comes from Mondaini et al. (2007), who investigated whether information affects side effects in men with prostatic hyperalgesia. These men were treated with finasteride for one year, a medication which is associated with side effects to do with sexual
function, including decreased libido, erectile dysfunction, and ejaculation disorders. Mondaini et al. (2007) randomised half of the men to be warned about these possible side effects and the other half to receive no such warning. At the end of the treatment period, reported occurrence of these side effects was much more frequent in men who were warned about possible side effects compared with those not warned. Specifically, 31% of men warned that finasteride could affect their sexual function reported erectile dysfunction, 24% reported decreased libido, and 16% reported ejaculation disorders compared with 10%, 8%, and 6%, respectively, in men not warned about this possibility.

In terms of the studies failing to find an effect of warnings about side effects, Howland et al. (1990) found no difference in side effects between general practice patients receiving erythromycin, an antibiotic, who were warned about possible side effects, including abdominal cramping and discomfort, and those not warned. Similarly, Myers and Calvert (1973) found no differences in side effects or withdrawal rates in patients with depression given amitriptyline and either warned or not warned about possible dizziness, dry mouth sweating, and constipation. Morris and Kanouse (1982) also failed to find an effect of warning patients given thiazide for hypertension about possible fatigue, skin rash or hives, dry mouth, dizziness, light-headedness, and headaches. Although, they did note that there was a trend towards those warned about side effects actually reporting more side effects.

One possible explanation for the studies failing to find an effect of providing information about side effects is that the studies involved various treatments with different possible side effects. In general, the placebo effect seems to affect some
conditions but not others (Evans, 2003). If so, some adverse effects may be more amenable to placebo-induced side effects than others and this may explain the differences in results across studies. Two studies conducted by Gibbs and her colleagues (1989a; 1989b) support this possibility. In both these studies, general practices in small towns in England were randomised, across towns, to provide their patients with prescription information leaflets or to provide no additional information. The information leaflets contained warnings about side effects relevant to the treatment being delivered. The first study (Gibbs et al., 1989a) involved patients given non-steroidal anti-inflammatory drugs, β-adrenoceptor antagonists, or bronchodilators and found no evidence that the information leaflet affected the side effects reported. The second study (Gibbs et al., 1989b), involving patients given penicillins, diuretics, or benzodiazepines, found that the information leaflet led to greater side effects in patients given benzodiazepines. Importantly, in each case the information led to increased knowledge of side effects, so these differences cannot be accounted for by differences in awareness of side effects. Instead, it may be that the treatment or symptoms related to benzodiazepines were amenable to placebo-induced side effects while the others were not. However, replication of the studies finding an effect of informing patients about treatment side effects would be required before these conflicting results could be attributed to whether or not the condition is placebo responsive.

There are also a number of general limitations to these studies. Firstly, they all involved active treatments. As discussed in Chapter 2, it is not yet known whether active treatment and placebo effects are additive. If they are not, then it may not be possible to isolate any placebo-induced side effects in studies involving active
treatment. In the studies that failed to find an effect of warning participants about side effects, for example, it could be that the side effects directly attributable to the active treatment might have produced a ceiling effect whereby no placebo-induced side effects could be detected. The only study that included a placebo treatment was that of Myers et al. (1987), but as mentioned above, they did not assess the effect of warning participants about side effects separately for those receiving active treatment and those receiving placebo. Secondly, all of these studies only considered the incidence of side effects and ignored their frequency and severity. This may not fully capture the variation in the side effects experienced by patients. For example, equal numbers of patients warned and not warned about side effects might experience one of these adverse symptoms, but the warning might increase the frequency or severity of the side effect.

These first two limitations suggest these studies were not maximally sensitive to placebo-induced side effects. However, a third limitation is that all of the side effects assessed in these studies were self reported. This means that the higher reporting of side effects found in some of these studies might simply have resulted from demand characteristics. This is a difficult limitation to overcome in these types of studies, because it is often not possible to objectively measure side effects, such as, headaches, dizziness, or dry mouth. Thus, while experimental studies on placebo-induced side effects provide at least some evidence that information about treatment side effects can lead to higher reporting of these side effects, they cannot determine whether this results from the placebo effect or some other process.
5.3 Conclusions

Studies showing that expectancies for adverse outcomes can produce negative placebo effects support the possibility that the placebo effect may contribute to treatment side effects. However, this type of evidence might only be considered indirect evidence for placebo-induced side effects as it involves expectancies regarding the primary action of the placebo, whereas placebo-induced side effects are secondary effects. Evidence that placebo groups in double-blind RCTs often report side effects is suggestive of placebo-induced side effects, but is limited in that these trials do not control for natural history. The fact that first time chemotherapy patients’ expectancies for nausea often predict their actual experience of post-chemotherapy nausea is equally suggestive. However, in the most methodologically rigorous study in this area, the results were ambiguous, with only one type of expectancy significantly predicting severe nausea, after history of nausea in some other settings was controlled for. Finally, at least four studies have shown that warning patients about side effects leads to a higher report of these side effects, but there are an equal number of studies that failed to find such an effect. Further, all of these studies involved active treatment, making it difficult to isolate the placebo effect, and they all only considered the incidence of side effects, rather than also examining their frequency and severity. Overall then, there seems to be evidence of a relationship between expectancies and side effects, but it remains unclear whether or to what extent the placebo effect actually contributes to treatment side effects.

With this in mind, the two following studies aimed to test for placebo-induced side effects by improving on some of the methodological limitations of the studies discussed in this chapter. In the first, I conducted a series of experiments to test
whether information about treatment side effects leads to greater side effects when only placebos were administered. To do this, I manipulated the side effect warnings given to otherwise healthy volunteers suffering from sleep difficulty who were told that they were taking part in a trial of a new medication designed to improve their sleep quality, but were actually given placebos. In the second study, I investigated whether first time chemotherapy patients’ pretreatment expectancies uniquely accounted for a significant proportion of variability in their posttreatment nausea, in a large sample (n>600) of cancer patients and after controlling for a number of other possibly confounding variables.
Chapter 6 – Information and Placebo-Induced Side Effects: Three Dummy Trials for Sleep Difficulty (Study 3)

6.1 Introduction

This study investigated whether providing information about possible treatment side effects leads to more side effects via the placebo effect. This involved a series of three dummy (placebo only) trials for sleep difficulty. Sleep difficulty was chosen because there is already some evidence for placebo effects in this area (e.g. Fratello et al., 2005; McCall, D'Agostino, & Dunn, 2003; Suetsugi et al., 2007) and it was expected to be relatively common in the student population from which participants were recruited, given prevalence of up to 35% in the general population (Stein, Belik, Jacobi, & Sareen, 2008). In these dummy trials, participants suffering from mild sleep difficulty were allocated to receive placebo treatment under the guise of a new medication for sleep difficulty or to a no treatment control group. The information participants received about bogus possible side effects was manipulated such that some participants were warned about side effects and others were not (Experiment 1) or such that the side effects participants were warned about were counterbalanced (Experiments 2 and 3). In the latter case this involved suggesting different sets of side effects to two groups of participants receiving placebo treatment.

The design of these experiments had a number of advantages over previous studies in this area. As discussed in Chapter 5, all experimental studies investigating placebo-induced side effects to date have involved active treatments (Gibbs et al., 1989a, 1989b; Howland et al., 1990; Mondaini et al., 2007; Morris & Kanouse, 1982; Myers & Calvert, 1973, 1984; Myers et al., 1987). If the active treatment produces
side effects in and of itself, then this may create a ceiling effect whereby any placebo-
induced side effects cannot be detected. Using placebo treatment avoids this
possibility and thereby increases sensitivity for detecting placebo-induced side effects.
The previous studies also only assessed the incidence of side effects as present or
absent. Such crude measures of assessment may be insufficient to fully capture the
impact of information on treatment side effects. For example, informing participants
about side effects might cause a placebo effect that increases the frequency or severity
of the symptom rather than increasing its occurrence. For this reason, the current
study also included assessment of the frequency and severity of each side effect.
Further, the inclusion of a no treatment group allowed for assessment of natural
history, in that the occurrence, frequency, and severity of the suggested side effects
could be observed in participants not receiving placebo treatment.

6.2 Experiment 1

Experiment 1 piloted the design for the current study. The main aim was to
determine whether the placebo treatment would elicit a placebo effect for sleep
difficulty: without such an effect it would be unrealistic to expect placebo-induced
side effects. A secondary aim was to determine whether warning participants about
side effects had any impact on the occurrence, frequency, or severity of the suggested
side effects. However, because only a small number of participants were involved in
this experiment, it was highly unlikely that differences in side effects, if any, would
reach statistical significance.
6.3 Method

6.3.1 Participants

Twenty-six first year (16 female, mean age=18.9, SD=1.3) psychology students from the University of Sydney experiencing mild difficulty sleeping participated in this study. Interested participants responded to an advertisement placed on ‘Experimtrix’, the School of Psychology, University of Sydney’s recruitment website for first year psychology students. Potential participants were excluded if they were under 18, had received treatment for sleep difficulty in the previous 3 months, were lactose intolerant, or were currently taking prescription medication other than the contraceptive pill. Participants received course credit for completing the study.

6.3.2 Design

Table 6.1 summarises the design of Experiment 1. The experiment employed a simple one factor, three level between-subjects design. One group received placebo treatment which they were told was a new herbal medication that would help them with their sleep difficulty. They also received an information sheet that suggested that the treatment may have three minor side effects: being more sleepy than usual when waking up, having a dry mouth, and mild headaches. A second group also received

Table 6.1. Design of Experiment 1. Participants were allocated to receive placebo treatment or no treatment (control). Half of those allocated to receive placebo treatment were warned about three possible side effects (suggestion) while the other half were not (no suggestion).

<table>
<thead>
<tr>
<th>Group</th>
<th>Treatment</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggestion</td>
<td>Placebo pills (lactose)</td>
<td>New treatment for sleep difficulty that may cause side effects, i.e. headaches, dry mouth, sleepy in the morning</td>
</tr>
<tr>
<td>(n=9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Suggestion</td>
<td>Placebo pills (lactose)</td>
<td>New treatment for sleep difficulty but no mention of side effects</td>
</tr>
<tr>
<td>(n=8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>No treatment</td>
<td>Controls in a trial of a new treatment for sleep difficulty</td>
</tr>
<tr>
<td>(n=9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
placebo treatment which they were told was a new herbal medication that would help them with their difficulty sleeping but received an information sheet that made no mention of side effects. Finally, a third group, the control group, did not receive treatment, but were told that they were acting as controls for a trial of a new sleep difficulty treatment. The dependent variables for sleep difficulty were global sleep difficulty, number of minutes taken to fall asleep (sleep latency), and total sleep time. The dependent variables for placebo-induced side effects were free and cued report of the occurrence, frequency, and severity of the suggested side effects.

6.3.3 Materials

**Placebo Pills:** The placebo pills were made from lactose and were red coated. They were prepared by the Faculty of Pharmacy, University of Sydney.

**Pittsburgh Sleep Difficulty Index (Amended; PSQI-A):** This is a standardised 19-item questionnaire developed by Buysse, Reynolds, Monk, Berman, and Kupfer (1989). It consists of seven components assessing subjective sleep difficulty, sleep latency, total sleep time, habitual sleep efficiency, sleep disturbances, use of medications, and daytime dysfunction over the previous month. Summing these 7 components provides a global sleep difficulty score ranging from 0-21 with higher scores indicating more sleep difficulty. Two amendments were made to this questionnaire. First, the component regarding the use of sleep medications was removed because part of the experimental manipulation involved administering treatment to some participants, but not others. This meant that total scores could range from 0-18 instead of 0-21. Secondly, because participants only received treatment for one week the questionnaire was amended such that participants rated their sleep difficulty over the previous week, rather than the previous month.
**Posttreatment Questionnaire (Appendix 4b):** This was a study-specific questionnaire given only to those who received the placebo treatment. It assessed side effects as well as other aspects of the treatment, such as, perceived efficacy and compliance. The side effect question asked participants to rate whether or not they had experienced any side effects during the treatment period and, if so, to specify what these side effects were as well as rate their severity. This provided their free report of side effects. The perceived efficacy question required participants to rate how effective they found the treatment was for reducing their sleep difficulty on a scale from 0, ‘Not at all’ to 10, ‘Very effective’. In terms of compliance, participants were asked whether or not they had taken all of the pills given to them. If they had not, they were asked to report on how many nights they had failed to take the pill and whether there was a reason for this.

**Adverse Symptom Checklist (Appendix 4c):** This was also a study-specific questionnaire. It assessed general adverse health symptoms and had the three target symptoms, sleepiness in the mornings, headaches, and dry mouth, embedded in it. This questionnaire required participants to rate 1) whether or not they had experienced the symptom in the previous week (occurrence), 2) the number of days on which they had the symptom (frequency), and 3) how severe the symptom had been (severity). The latter was on an 11-point Likert-type scale anchored by ‘Not at all severe’ to ‘Extremely severe’. The checklist reflected participants’ cued report of side effects and was given to all participants.
6.3.4 Procedure

Participants attended an initial session lasting approximately 20min in which they were given the information sheet and consent form. The information sheet (Appendix 4a) included a background section which described the experiment as a trial of a new herbal medication, labelled SX3752, believed to be effective for reducing sleep difficulty and stated that they may be allocated to receive the treatment or to a group that would not receive treatment. Half of the information sheets contained a warning about possible side effects under a section titled ‘Risks’. This warning stated:

“Some people who take SX3752 report some minor side effects. These may include feeling more sleepy than usual when you awaken, having a dry mouth, and mild headaches. If you do experience these side effects, they will go away completely when you stop taking SX3752”

The remaining information sheets did not contain this section and made no mention of possible side effects. After informed consent was gained participants were asked to provide some demographic information and complete a questionnaire about their sleeping habits over the past week (PSQI-A). This constituted their baseline sleep difficulty. Participants were then informed about whether they would be receiving the treatment or not. Those who were in one of the two treatment groups were given 7 placebo pills (lactose) and told to take one each night for the next week about 15 minutes before going to bed. Those in the no treatment group were told that they were acting as controls and would not be receiving treatment. Participants returned one week later to complete the follow-up assessment which lasted approximately 15min. In this session, participants were asked to complete the PSQI-A which assessed their level of sleep difficulty over the treatment period. Those who received treatment were
then given the posttreatment questionnaire which assessed side effects via free report, perceived efficacy of the treatment, and compliance. All participants were then given the adverse symptom checklist which had the three suggested side effects embedded in it. At the end of the session participants were debriefed and given information about the ‘Quarter of an Hour Rule’ (Bootzin, Epstein, & Wood, 1991), which is a behavioural strategy that can be used to reduce sleep difficulty. These procedures were approved by the University of Sydney’s Human Research Ethics Committee.

6.3.5 Statistical Analysis

Two orthogonal contrasts tested whether placebo treatment reduced sleep difficulty compared with no treatment at the follow-up session and whether there were differences in sleep difficulty between the suggestion and no suggestion groups. This was done on posttreatment scores for global sleep difficulty, sleep latency, and total sleep time, controlling for baseline scores. Only a few participants freely reported side effects and the symptoms varied from one participant to another making it impractical to compare frequency and severity. As a result, only the occurrence of any adverse effects was considered when comparing freely reported side effects and this was done using a Chi-square test of independence. For cued report of side effects, Chi-square tests of independence assessed the occurrence of the suggested side effects across all groups. Orthogonal contrasts were then used to assess the frequency and severity of each suggested side effect in those reporting these symptoms for placebo treatment versus no treatment and the suggestion versus no suggestion groups. All statistical analyses were conducted using SPSS software (version 15; SPSS Inc, Chicago, Ill) and results were considered significant when $p<.05$. 
6.4 Results and Discussion

6.4.1 Sleep Difficulty

Figure 6.1 displays mean global sleep difficulty before and after treatment across groups. Global sleep difficulty decreased by an average of 4.7 (SD=2.5) points for those in the suggestion group, 3.0 (SD=2.0) points for those in the no suggestion group, and 2.1 (SD=3.7) points for those in the control group. Contrast analysis revealed that posttreatment global sleep difficulty was significantly lower for those who received placebo treatment compared with no treatment, after controlling for baseline scores, $F(1,22)=5.29, p=.03$, but that there was no significant difference in global sleep difficulty among the two treatment groups, $F<1$.

![Global Sleep Difficulty](image)

*Figure 6.1. Experiment 1. Mean (+SE) global sleep difficulty across treatment group. Placebo treatment significantly reduced sleep difficulty but there was no significant difference in sleep quality between the suggestion and no suggestion groups.*

Mean sleep latency and total sleep time for each group are shown in Figure 6.2. There was a tendency for a greater reduction in sleep latency for those receiving
placebo treatment compared with the control participants, which was only marginally non-significant after controlling for baseline sleep latency, \( F(1,22)=4.26, p=.051 \). There was also a tendency for increased total sleep time for participants receiving placebo compared with those receiving no treatment, however, this was also non-significant after controlling for baseline total sleep time, \( F<1 \). There were no differences between the suggestion and no suggestion groups on either of these measures of sleep difficulty, both \( F<1 \). This pattern of results suggests a placebo effect for global sleep difficulty and tendencies towards the same for sleep latency and total sleep time, but that this placebo effect was unaffected by whether or not participants were warned about side effects.

**Figure 6.2.** Experiment 1. Mean (+SE) sleep latency (A) and total sleep time across groups. There were non-significant tendencies towards better sleep quality for those receiving placebo treatment compared with those receiving no treatment. There were no significant differences between the suggestion and no suggestion groups.

### 6.4.2 Side effects – Free report

Six out of the seventeen (35%) participants who received placebo treatment freely reported experiencing at least one side effect. The rate of these side effects was almost identical with three participant reporting at least one side effect in the
suggestion group (33%) and three participants in the no suggestion group (38%), 
\( \chi^2<1 \). In the suggestion group two of the participants reported side effects consistent with the warning: one participant reported all three suggested side effects (tiredness in the morning, headaches, and dry mouth) while the second reported only dry mouth. The third participant in this group reported experiencing stomach cramps on the last night of taking the tablet. The side effects reported by two of the participants in the no suggestion group were also consistent with the target side effects even though these participants received no warning: they were, tiredness during the day and headaches. The third participant in the no suggestion group reported experiencing difficulty concentrating. This suggests that participants receiving placebo treatment report experiencing side effects, but that the warning they received that had no affect on the type of these side effects.

6.4.3 Side effects – Cued report

Figure 6.3 displays the occurrence of target symptoms as a function of treatment group. A higher proportion of participants in the suggestion group reported sleepiness in the morning and headaches than the no suggestion group while the opposite was true for dry mouth. The occurrence of these symptoms was quite high in the control group with at least 50% of these participants reporting each of the three target symptoms. Chi-square tests, however, revealed that there were no significant differences in occurrence of these three target symptoms between groups, highest \( \chi^2(df=2, n=26)=1.65, p=.44 \).
Figure 6.3. Experiment 1. Occurrence (%) of target symptoms across treatment groups. No differences were significant.

As shown in Figure 6.4, sleepiness in the mornings and dry mouth were rated as being slightly more frequent in the suggestion group compared with the no suggestion group, but these differences were not significant, both $F<1$. For headaches, however, the no suggestion group actually reported this side effect as more frequent than the suggestion group, although this was marginally non-significant, $F(1,8)=4.64$, $p=.06$. Frequency of side effects in the control group was lower than both treatment groups, with the exception of headaches in the suggestion group. These tendencies towards less frequent side effects in the control group were not significant, highest $F(1,14)=3.04$, $p=.11$. 
**Side Effects - Frequency**

![Graph showing frequency of side effects across treatment groups]

**Target Symptoms**

- Sleepy Mornings
- Headaches
- Dry Mouth

**Figure 6.4:** Experiment 1. Mean (+SE) frequency in days each target symptom occurred across treatment group. No differences were significant.

Figure 6.5 shows a small tendency for those in the suggestion group to report greater symptom severity than those in the no suggestion group, however none of these differences were statistically significant, all $F<1$. The mean severity of the three target symptoms did not differ significantly between those receiving placebo treatment and those in the control group, all $F<1$. Taken together, the findings for occurrence, frequency, and severity of the target side effects indicate that the warning had minimal, if any, affect on these adverse symptoms.
6.4.4 Summary

Even with this small sample there was clear evidence of a placebo effect for sleep difficulty. Participants who received a placebo pill but were told that they were receiving an active medication reported significantly less global sleep difficulty at the end of the treatment period compared with participants who received no treatment. This was supported by tendencies towards reduced sleep latency and increased total sleep time in the placebo treatment groups compared with the control group. In relation to side effects, the suggestion that the treatment may result in increased tiredness in the morning, mild headaches, and dry mouth appeared to have minimal, if any, effect on the occurrence, frequency, or severity of these symptoms, however. It was interesting that one third of the participants receiving placebo treatment reported at least one side effect via free report, regardless of whether they were warned about side effects or not. This may provide some evidence for placebo-induced side effects.
as these participants may have expected and therefore experienced adverse symptoms as a result of these expectancies. On the other hand, these participants may have misattributed the random occurrence of these adverse symptoms during the experiment to the treatment. The fact that the control group, who never received treatment, often reported similar occurrence, frequency, and severity of the target side effects supports this possibility. However, this group had the highest level of sleep difficulty, which may have contributed to their high level of reported adverse symptoms.

6.5 Experiment 2

Overall, the results obtained in Experiment 1 were promising. They demonstrated that the design was sufficient to produce a placebo effect for sleep difficulty. The placebo treatment led to some side effects being reported via free report, but whether these were placebo-induced side effects or were simply misattributed could not be determined. There were some small tendencies towards increased occurrence, frequency, and severity of the target side effects in the suggestion group compared with the no suggestion group. Importantly, the differing patterns of results for side effect occurrence, frequency, and severity, support the use of more sophisticated analysis of side effects than those used in previous studies (e.g. Gibbs et al., 1989a, 1989b; Howland et al., 1990; Mondaini et al., 2007; Morris & Kanouse, 1982; Myers & Calvert, 1973, 1984; Myers et al., 1987).

A number of potential limitations were revealed during this pilot experiment. Some participants in the treatment groups questioned whether they had received a placebo at the end of the study period. This was surprising because participants were
told that they would receive either active treatment or no treatment, with no mention of placebos. This doubt regarding type of treatment might have weakened expectancies for reduced sleep difficulty and/or side effects and may, therefore, have detracted from the placebo effect for either of these outcomes. A second limitation was the high level of target symptoms in the control group, as noted above. This made it impossible to determine whether the side effects freely reported in the treatment groups occurred as a result of the placebo effect or of misattribution. Finally, the side effect manipulation may not have been strong enough to induce expectancies for these adverse effects. That is, participants in the suggestion group may not have attended to the warning contained in the information sheet sufficiently to cause an expectation for these adverse effects. If so, placebo-induced side effects would be unlikely.

Experiment 2 employed a similar design to Experiment 1, but a number of changes were made in an attempt to increase its sensitivity to both the placebo effect and placebo-induced side effects. The most significant of these was counterbalancing the side effects which the suggestion group were warned about. Participants in the suggestion group were informed that they might experience drowsiness, dry mouth, and nausea (Set A) or dizziness, blurred vision, and sore eyes (Set B). This provided a more sensitive measure of the influence of information on side effects because it allowed for a within-subjects comparison of suggested versus non-suggested side effects. Participants in the suggestion group also received a verbal warning about these possible, but bogus, side effects when they were given their treatment, in addition to the written warning contained in the information sheet. This was done to increase attention to the side effect warning. A question was also added to the end of the posttreatment questionnaire asking participants in the treatment groups whether
they remembered being warned about side effects or not, and, if so, which side effects they had been warned about in order to check the side effect manipulation. Finally, the information that the treatment was a new herbal medication was removed. Instead participants were simply told that the trial was for a new medication designed to reduce their sleep difficulty. This was aimed at increasing expectancies for both improved sleep quality and treatment side effects as some participants might expect herbal treatments to be less effective, or at least less toxic, than other types of treatment.

6.6 Methods

Except where otherwise stated, the methods used in Experiment 2 were the same as Experiment 1.

6.6.1 Participants

Participants were 57 (mean age=20.1, SD=3.9, female=44) first year psychology students from the University of Sydney who reported difficulty sleeping. Recruitment and eligibility criteria were identical to Experiment 1.

6.6.2 Design

The side effects which the suggestion group (n=18) were warned about were now counterbalanced. This meant that half of the participants in the suggestion group were warned about drowsiness, dry mouth, and nausea (Set A; Appendix 4d) while the other half were warned about dizziness, blurred vision, and sore eyes (Set B; Appendix 4e). Allocation to Set A or Set B side effects was random. As with Experiment 1, the no suggestion group (n=21) received placebo treatment but were not warned about side effects and the control group (n=18) did not receive treatment.
6.6.3 Procedure

All participants were told that they were taking part in an open trial of a new medication for sleep difficulty that did not contain a placebo group. Participants in the suggestion group received a verbal warning about either Set A or Set B side effects when they received their medication in addition to the warning contained in the information sheet.

6.6.4 Materials

*Adverse Symptom Checklist (Appendix 4f)*: The version described in Experiment 1 was altered to include all six side effects (Sets A and B) plus some other non-suggested adverse effects, e.g. restlessness.

*Posttreatment Questionnaire (Appendix 4g)*: This contained the extra question asking participants to rate whether they remembered being warned about side effects, and if so, to list these. This question occurred at the end of the questionnaire.

6.6.5 Statistical Analysis

Statistical analysis for the placebo effect for sleep difficulty and free report of side effects was identical to Experiment 1. The counterbalancing of side effects allowed for both within- and between-groups analysis for cued report of side effects in the suggestion group. For the within-subjects analysis, the total number of target and non-target side effects and their mean frequency and severity were calculated for each participant. Paired samples t-tests were then used to determine whether there were differences in the reporting of target versus non-target side effects. For the
between-subjects analysis, the occurrence, frequency, and severity of each adverse symptom were compared when they were suggested with when they were not suggested via Chi-square tests of independence and independent samples $t$-tests.

### 6.7 Results and discussion

#### 6.7.1 Sleep Difficulty

Figure 6.6 displays global sleep difficulty before and after treatment for each group. Sleep difficulty decreased by an average of 3.1 (SD=3.0) points in the suggestion group, 3.0 (SD=1.6) points in the no suggestion group, and 0.8 (SD=2.1) points in the control group. Controlling for baseline sleep quality, posttreatment sleep difficulty was significantly lower in those who received treatment compared with controls, $F(1,53)=10.6$, $p<.01$, but again there was no significant difference in global sleep quality between the treatment groups, $F<1$.

![Global Sleep Difficulty](image)

*Figure 6.6: Experiment 2. Mean (+SE) global sleep difficulty before and after treatment for each group. Placebo treatment significantly reduced sleep difficulty compared with no treatment. There was no significant difference between the two treatment groups (suggestion vs no suggestion).*
This pattern of results was the same for total sleep time, as shown in Figure 6.7B. Mean total sleep time increased by 33.5 (SD=56.3) minutes in the suggestion group and 22.9 (SD=40.1) minutes in the no suggestion group but only increased by 1.7 (SD=90.0) minutes in the control group. The increase in total sleep time for participants receiving placebo treatment was significantly higher than for those in the control group, after controlling for baseline total sleep time, $F(1,53)=4.4, p=.04$.

There was, however, no significant difference between the two treatment groups, $F<1$.

As can be seen in Figure 6.7A, there were no significant differences between groups for sleep latency after controlling for baseline scores, highest $F(1,53)=1.94, p=.17$.

The improved sleep quality in participants allocated to receive placebo treatment for global sleep difficulty and total sleep time was consistent with the findings of Experiment 1 and provides further evidence for a placebo effect for sleep difficulty.

---

*Figure 6.7. Experiment 2. Mean (+SE) sleep latency (A) total sleep time (B) and by treatment group. The only significant difference was that total sleep time was greater in those that received placebo treatment compared with controls.*
6.7.2 Side effects – Free report

Overall 12 out of 39 participants who received placebo treatment reported experiencing at least one side effect via free report. These were approximately equally distributed between the suggestion and no suggestion groups, 33.3% and 28.6%, respectively, $\chi^2<1$. The side effects reported in the suggestion group were not related to the warning. In fact, only two of these participants reported a target side effect and of these two, one also reported a non-target side effect. Combining the two treatment groups, the most common side effects reported were headaches and drowsiness with each being reported four times. Interestingly, one participant claimed to have vomited in response to the pill on her final night of treatment. Thus, as in Experiment 1, participants receiving placebo treatment for sleep difficulty freely reported side effects in response to their treatment, but these were not related to whether or not they had been warned about side effects.

6.7.3 Side effects – Cued report

Figure 6.8 displays average occurrence, frequency, and severity of target versus non-target symptoms reported by the suggestion group via cued report. Although there was a slight tendency for more target symptoms to be reported by the suggestion group than non-target symptoms and for the target symptoms to be more frequent and severe, paired $t$-tests revealed that none of these differences were significant, all $t<1$. In general, the mean occurrence, frequency, and severity of both target and non-target symptoms were quite low.
Figure 6.8: Experiment 2. Mean (+SE) occurrence (total number of symptoms, A), frequency (B), and severity (C) of target symptoms versus non-target symptoms reported by the suggestion group. No differences were statistically significant.

As shown in Figure 6.9, between-subjects analysis of each of the six target symptoms when suggested versus not suggested showed a similar pattern to the within-subjects analysis above. That is, there was a general tendency for symptoms to have greater occurrence, frequency, and severity when they were suggested than when they were not suggested. The only exceptions to these were occurrence, frequency, and severity of dizziness and occurrence of drowsiness. Chi-square tests of independence and independent samples t-tests revealed that none of these differences were statistically significant, all $\chi^2<1$ or $t<1$. As can also be seen in Figure 6.9, the occurrence, frequency, and severity of each of these side effects were relatively high in the control group. Thus, these results indicate a similar pattern to those of Experiment 1, in that there were small tendencies for increased side effects when suggested, but that none of these differences were statistically significant.
Figure 6.9. Experiment 2. Occurrence (A) and mean (+SE) frequency (B) and severity (C) of each symptom in the suggestion group when suggested (target), in the suggestion group when not suggested (non-target), and in the control group (control). No differences were statistically significant.
6.7.4 Recall of side effect warning

Recall of the side effect warning was poor. Only one participant (6%) in the suggestion group correctly recalled all three of the side effects contained in the warning and, as shown in Figure 6.10, almost two thirds of these participants could only correctly recall one or less of the side effects that they had been warned about. One third of these participants also incorrectly reported being warned about a symptom that they had, in fact, not been warned about.

![Recall of Suggested Side Effects](image)

*Figure 6.10. Experiment 2. Correct recall of side effects contained in the warning given to participants in the suggestion group.*

6.7.5 Summary

Experiment 2 demonstrated further evidence for a placebo effect for sleep difficulty. Participants who received placebo treatment showed significantly lower levels of global sleep difficulty and had longer total sleep time than those who received no treatment. The warning about side effects did not significantly increase occurrence, frequency, or severity of side effects via free or cued report. There were tendencies in this direction, but they were only small. It was again interesting to note,
however, that approximately 30% of all participants who received the placebo
treatment reported at least one side effect when asked. Perhaps most importantly, it
was clear that the majority of participants could only remember being warned about
one side effect or less at the end of the treatment period. This suggests that
participants either failed to attend to the warning about side effects when given or that
substantial forgetting occurred during the treatment period. Given such poor recall, it
is not surprising that the warning had no impact on participants’ report of side effects.

6.8 Experiment 3

Poor recall of the warning was a major and unexpected limitation to
Experiment 2. To develop expectancies for suggested side effects participants must
first attend to and retain the information contained in the warning. The poor recall was
unexpected because participants received an extra verbal warning in the experiment.
One possible explanation is that participants received a large amount of information in
a relatively small amount of time. The fact that participants remembered being
warned about possible side effects, for the most part, but forgot which specific side
effects they had been warned about supports this. Although the participants were
given the information sheet to take home with them, it is unlikely that they would
have re-read the information during the treatment period.

Experiment 3 addressed this limitation by reducing the warning to contain
only one side effect, rather than three. This was intended to reduce the memory load
on participants regarding the warning and, thereby, to increase their expectancies for
the suggested side effect. Three other modifications were made to the design of
Experiment 3. The first was that the side effects participants were warned about were
changed to either increase in appetite or decrease in appetite. The advantage of this
was that these symptoms are symmetrical, in that an increase in appetite precludes a decrease in appetite and vice versa. Unfortunately, there were no data on occurrence, frequency, and severity of change in appetite likely in the control group from the previous experiments, as this was not included on either of the adverse health checklists. The second was that the no suggestion group was omitted. This was done because the counterbalancing of the side effect warning allowed for a within-groups comparison of suggested versus non-suggested side effects in the suggestion group. The control group was maintained, however, as it was important to confirm the placebo effect for sleep difficulty and to determine the extent to which increased or decreased appetite was reported by participants not receiving treatment. Finally, the colour of the placebo pills was changed to blue, rather than red. This was based on evidence indicating that blue pills are associated with greater hypnotic effects than red pills (Blackwell, Bloomfield, & Buncher, 1972; Buckalew & Coffield, 1982; de Craen, Roos, Leonard de Vries, & Kleijnen, 1996) and that increasing the placebo effect for sleep difficulty might increase expectancies for treatment side effects.

6.9 Methods

Except where otherwise stated, the methods used in Experiment 2 were the same as Experiment 2.

6.9.1 Design

The no suggestion group was omitted from this study. Therefore, participants were either allocated to receive placebo treatment under the guise of new medication aimed at reducing their sleep difficulty or to a no treatment control group. Half of the participants received both a written and verbal warning that the treatment might cause
increased appetite (Appendix 4h), while the other half received a written and verbal warning that the treatment might cause a decrease in appetite (Appendix 4i). This occurred whether or not participants received treatment so as to avoid any potential sampling bias which may have arisen if the warning deterred participants from joining the study. Therefore, half of the participants in the control group were also given an information sheet that warned of possible increased appetite in response to the bogus treatment, while the other half received an information sheet that warned of possible decreased appetite. Participants were allocated to the suggestion or control group on a 3:2 ratio to ensure a sufficient sample size in the experimental group, which resulted in 31 participants in the suggestion group and 17 participants in the control group.

6.9.2 Participants

Participants were 48 (mean age=20.1, SD=5.1, female=37) first year psychology students from the University of Sydney who reported difficulty sleeping. Eligibility criteria were identical to Experiment 2.

6.9.3 Procedure

The procedure used was the same as Experiment 2, with the exception that participants were either warned about increase in appetite or decrease in appetite.

6.9.4 Materials

Placebo Pills: The placebo pills were made from lactose and were blue coated. They were prepared by the Faculty of Pharmacy, University of Sydney.

Adverse symptom checklist (Appendix 4k): The version used in Experiment 2 was altered to include a question asking whether the participant had experienced a change in appetite, and if so to describe this and to rate its severity. The direction of
the change in appetite was derived from the description participants provided. Due to experimenter error, the question regarding the frequency of change in appetite was unintentionally omitted, so only the occurrence, direction, and severity of change in appetite were assessed.

6.9.5 Statistical Analysis

ANCOVA was used to test whether participants receiving placebo treatment had lower levels of sleep difficulty at the follow-up compared with those receiving no treatment and after controlling for baseline levels. For side effects via cued report, a nonparametric measure was calculated which indicated appetite changes as follows: 1) consistent with the warning, 2) inconsistent with the warning, and 3) no change in appetite. A Chi-square test of independence was used to determine whether the pattern of changes in appetite differed among those in the suggestion group and those in the control group. A Chi-square test for goodness of fit assessed whether there was an overall effect of the warning, regardless of treatment, by comparing the proportion of consistent changes in appetite with those that were inconsistent in all participants reporting a change in appetite. An independent samples t-test was used to compare the severity of consistent changes in appetite across the suggestion and control groups. Two participants in the control group failed to answer the question regarding changes in appetite so they were omitted from the analysis of side effects based on cued report. No statistical analysis was conducted on free report of side effects because only four participants reported side effects in response to this question.
6.10 Results and discussion

6.10.1 Sleep difficulty

The pattern of results for sleep difficulty replicated the placebo effect found in Experiment 2. As shown in Figure 6.11, controlling for baseline scores, placebo treatment significantly reduced global sleep difficulty by 1.6 points and increased total sleep time by more than no treatment and increased to 26min compared with the control group who received no treatment, $F(2,45)=4.69, p=.04$ and $F(2,45)=8.02, p<.01$, respectively. There was again no significant difference in sleep latency between the suggestion and control groups, $F(2,45)=2.01, p=.16$. Thus, there was further evidence for a placebo effect for sleep difficulty.

![Figure 6.11. Experiment 3. Mean (+SE) global sleep difficulty (A) and total sleep time before and after treatment for the suggestion and control group. In both cases, placebo treatment led to significantly better posttreatment sleep quality.](image)

6.10.2 Side effects – Free report

Only four (13%) of the participants receiving placebo treatment freely reported a side effect. Two of these participants reported a change in appetite and on both occasions these were consistent with the warning these participants had received. The third participant reported migraines and drowsiness, while the fourth reported...
difficulty waking up and strange dreams. The proportion of participants freely reporting side effects was clearly lower in this experiment than in the previous two experiments (approx 30%).

6.10.3 Side effects – Cued report

Figure 6.12 displays the proportion of participants reporting consistent, inconsistent, or no change in appetite for the suggestion group and the control group. Eight (26%) participants in the suggestion group reported a change in appetite and these were all in the direction of the warning they had received. This suggested a strong impact of the warning on appetite. However, four participants in the control group also reported a change in appetite and three (20% overall) of these were consistent with the warning contained in the information sheet that these participants received. A Chi-square test of independence revealed no significant difference in the pattern of changes in appetite for the suggestion group compared with the control group, $\chi^2(df=2, n=46)=2.21, p=.33$. A post hoc Chi-square test for goodness of fit showed that, regardless of treatment condition, the direction of changes in appetite was significantly associated with the warning participants had received, $\chi^2(df=1, n=12)=8.33, p<.01$. Taken together, this implied that there was an overall effect of suggestion whereby participants reporting a change in appetite did so in the direction which they had been warned about, but that this was not significantly related to whether or not they actually received the placebo treatment. This finding was rather unexpected and is discussed in more detail below. Figure 6.13 displays severity ratings of the changes in appetite for the suggestion and control groups. There were no significant differences between these two groups, $t<1$. 
Change in Appetite - Direction

Figure 6.12. Experiment 3. Proportion and direction of changes in appetite for participants in the suggestion and control groups. No differences were statistically significant.

Change in Appetite - Severity

Figure 6.13. Experiment 3. Mean (+SE) severity of changes in appetite for participants in the suggestion and control groups. There was no statistically significant difference.

6.10.4 Recall of the side effect warning

Recall of the side effect warning was near perfect. All but one participant in the suggestion group recalled being warned about side effects. Further, all of the participants who remembered being warned about side effects correctly identified the
direction of change in appetite that had been suggested. This indicates an important improvement over Experiment 2, where almost two thirds of the participants could only remember being warned about one or less side effects from a possible three.

6.10.5 Summary

Experiment 3 replicated the placebo effect for sleep difficulty found in Experiment 2, with those receiving placebo treatment reporting significantly lower global sleep difficulty and increased total sleep time posttreatment compared with control participants. There was a smaller proportion of side effects freely reported in this experiment (13%) compared with the two previous experiments (about 30%). The findings for changes in appetite when cued were rather odd, in that there seemed to be an overall effect of warning participants about possible changes in their appetite, but that this was unrelated to whether or not the participants actually received placebo treatment. The warning about side effects was recalled much better in this experiment than in Experiment 2, with virtually all of the participants in the suggestion group remembering the side effect they had been warned about.

6.11 General Discussion

There was consistent evidence for a placebo effect for sleep difficulty in all three experiments. In Experiments 1, 2 and 3, placebo treatment led to significantly lower global sleep difficulty than no treatment and in Experiments 2 and 3 placebo treatment also significantly increased total sleep time. This supports previous evidence for placebo effects in sleep difficulty (e.g. Fratello et al., 2005; McCall et al., 2003; Suetsugi et al., 2007). It is worth noting that the control groups consistently demonstrated reductions in their sleep difficulty, even though they did not receive
treatment. *Post hoc* analysis of pooled data from the control groups in each experiment demonstrated that these reductions were significant for global sleep difficulty and sleep latency. It is not clear why this was the case. It may reflect 1) an initial reporting bias during, 2) that taking part in the trial itself improved their sleep quality, or 3) that there was regression to the mean. Regardless of the reason, it re-emphasises the need to control for natural history when investigating the placebo effect.

In relation to placebo-induced side effects, the first two experiments provided no statistically significant evidence that warning participants about possible side effects actually influenced their reports of these adverse symptoms. However, this is perhaps unsurprising given the poor recall of the suggested side effects found in Experiment 2. In Experiment 3, in which almost all participants correctly recalled the side effect they had been warned about, 26% of participants who received placebo treatment reported the suggested side effect, either increase or decrease in appetite. Further, no participants in this group reported a change in appetite that was inconsistent with the warning. Although this result may suggest that the warning affected side effects, this pattern was not significantly different to the control group, in which changes in appetite also generally appeared consistent with the warning, even though these participants did not receive treatment. As such, this implies an overall effect of warning participants about potential side effects that was independent of whether participants actually received treatment or not. However, the control group only consisted of seventeen participants and only four of these participants reported a change in appetite, so it is difficult to determine whether this result would hold in a larger sample.
Despite the lack of convincing evidence for placebo-induced side effects caused by the warning, a number of participants receiving placebo treatment reported experiencing at least one side effect in response to their treatment. In the first two experiments approximately 30% of participants receiving placebo treatment reported at least one side effect, whereas only 13% did in the final experiment. This may provide evidence for placebo-induced side effects, if participants entered the experiment with preconceived expectancies for side effects. The difference in rates of freely reported side effects in Experiments 1 and 2 versus Experiment 3 is potentially interesting. The two major differences between these experiments were that the former involved a warning about three possible side effects and that the placebo pills were red, while the latter involved a warning about only one possible side effect and the placebo pills were blue. Perhaps being warned about a greater number of side effects created more generalised expectancies for adverse outcomes than only being warned about one side effect. Equally plausible is that red pills might be more readily associated with toxicity compared with blue pills and this may have led to stronger expectancies for side effects. Of course, caution is required here as the high rate of adverse symptoms in the control group implies that the side effects reported by the treatment groups might have simply been misattributed to the treatment, rather than actually being caused by the placebo effect.

Overall then, there was generally little evidence for placebo-induced side effects, which is consistent with previous studies that have failed to find evidence that information about side effects increases their occurrence (Gibbs et al., 1989a; Howland et al., 1990; Morris & Kanouse, 1982; Myers & Calvert, 1973). However,
given the small but consistent tendencies towards increased side effects in the suggestion groups, it might be the case that placebo-induced side effects do exist, but that they are relatively weak effects. As discussed in the previous chapter, placebo-induced side effects might be considered secondary placebo effects because they are responses made to the treatment other than those for which it has been administered and that are not attributable to the inherent properties of the treatment itself. As such, participants might develop much stronger expectancies for the primary outcome of their treatment, in this case reduced sleep difficulty, than they do for side effects. This certainly seems plausible given the strong evidence for a placebo effect for sleep difficulty in the current study. Further, it may explain why there is more convincing evidence for negative placebo effects than there is for placebo-induced side effects (see Chapter 5).

There are at least two important implications of this possibility, both of which are discussed in more detail in Chapter 8. Firstly, if placebo-induced side effects do exist but are only weak, then detecting these effects will require very large sample sizes, which, despite employing a highly sensitive design, the current study lacked. Secondly, the fact that participants appear to benefit more from the suggestion of positive outcomes than the suggestion of adverse outcomes implies that providing information about a treatment, its aims, likely efficacy, and side effects will produce more benefit via the placebo effect than it will harm. This means that, in addition to allowing for increased patient autonomy, informed consent might improve outcomes via the placebo effect substantially more than it increases adverse side effects.
The current study also highlights a general limitation to experimental studies assessing the influence of warning participants about side effects. That is, these types of studies do not directly assess expectancies for side effects. Such assessment was deliberately avoided in the current study because questioning participants about their expected side effects might influence their expectancies or it might have undermined the credibility of the bogus trial and hinted towards the true purpose of the study. As a result, it was impossible to determine whether participants actually expected to experience side effects as result of the warning, beyond recall of the side effects listed. If no such expectancies were elicited, then no placebo-induced side effects could be expected. Therefore, while experimental studies involving manipulating information about treatment side effects are best able to determine whether this has a causal impact on actual side effects, they are entirely reliant on the side effect warning producing expectancies for side effects, which for the above reasons is difficult to check.

In an attempt to explore the possibility that placebo-induced side effects do exist but only have small effects, the following study investigated whether expectancies contribute to treatment side effects in a large sample of first time chemotherapy patients. An advantage of this was that the relationship between expectancy and side effects could be assessed directly.
Chapter 7 – Expectancy and Posttreatment Nausea in First Time Chemotherapy Patients (Study 4)

7.1 Introduction

This study examined the relationship between expectancy and nausea in a large sample of first time chemotherapy patients in a trial comparing antiemetic regimens for delayed nausea. Although methods for the prevention and control of emesis have improved greatly, nausea continues to be a significant burden to patients undergoing chemotherapy, with the vast majority of these patients experiencing nausea at some point during their chemotherapy treatment (Roscoe et al., 2000). As discussed in Chapter 5, a number of studies have shown that stronger pretreatment expectancies for nausea predict greater posttreatment nausea (Haut et al., 1991; Jacobsen et al., 1988; Montgomery & Bovbjerg, 2000; Olver et al., 2005; Rhodes et al., 1995; Roscoe et al., 2004; Roscoe et al., 2000; Shelke et al., 2008). While such studies are suggestive of placebo-induced side effects, the majority of these studies have methodological flaws limiting the conclusions which can be drawn from them. For example, many did not adequately control for other potentially confounding variables, particularly history of nausea, and/or used statistical techniques that may have overestimated the role of expectancy. This makes it difficult to determine whether expectancy actually contributes to post-chemotherapy nausea or whether the two are simply correlated.

In the only study that avoided all of the above limitations (Roscoe et al., 2004) the results were equivocal. Here, patients’ retrospective ratings of their expectancies before speaking to their oncologist significantly predicted the occurrence of severe
nausea in the 5 days following their first infusion, but not average nausea. Further, despite measuring susceptibility to motion sickness, which has been shown to relate to post-chemotherapy nausea (Jacobsen et al., 1988; Morrow, 1985; Roscoe et al., 2000), the authors did not control for this potentially confounding variable.

The current analysis aimed to overcome these limitations in order to determine whether expectancy contributes to post-chemotherapy nausea and, thereby, provide evidence for placebo-induced side effects. To do this, I used hierarchical regression to evaluate whether first time chemotherapy patients’ pretreatment expectancies were significantly associated with their posttreatment nausea, over and above other possible factors, such as age, gender, diagnosis, susceptibility to motion sickness, and pretreatment quality of life (QoL). Importantly, the large sample size meant that the study was well powered to detect even small effects. Further, there was a large sub-sample of previously pregnant women, which enabled me to assess whether including nausea during pregnancy as a covariate influenced the relationship between expectancy and post-chemotherapy nausea.

An interesting possibility not addressed in previous studies is that the relationship between expectancy and post-chemotherapy nausea may be non-linear. If it is not linear, then a certain level of expectancy may be required to produce an effect. For instance, patients with particularly high levels of expectancy might experience more nausea than those with moderate and low levels of expectancy with little difference between the lower levels themselves. Conversely, having very low expectancies for nausea might have a protective effect and result in less nausea compared with those who have moderate or high levels of expectancy with little
difference between the higher levels. This possibility was addressed by categorising patients into groups according to their level of expectancy and investigating whether a particular level of expectancy either heightened or protected against post-chemotherapy nausea.

Predictors of patients’ expectancies for nausea were also examined in order to determine whether a history of nausea in other areas is associated with expectancies for post-chemotherapy nausea. In the only previous study addressing this, Montgomery and Bovbjerg (2003) found that patients’ lifetime history of nausea, including susceptibility to motion sickness and nausea during pregnancy, were unrelated to their expectancies for post-chemotherapy nausea. However, this analysis was based on only 31 patients. Finally, the inclusion of a measure of quality of life allowed me to assess the relationship between post-chemotherapy nausea and quality of life. Nausea is often rated as one of the most severe and debilitating side effects of chemotherapy (Carelle et al., 2002; Griffin et al., 1996; Klastersky, Schimpff, & Senn, 1999) and has been shown to reduce quality of life (Ballatori & Roila, 2003; Ballatori et al., 2007; Cohen et al., 2007; Lindley, Hirsch, O'Neill, Transau, & et al., 1992; Osoba et al., 1997). However, no study to date has simultaneously investigated expectancies, post-chemotherapy nausea, and quality of life in cancer patients.

7.2 Methods

7.2.1 Participants

Participants were first time chemotherapy patients taking part in a multicentre trial comparing antiemetic regimens for the treatment of delayed nausea. They were enrolled from 18 private practice oncology groups in the USA between June 12,
2001 and June 11, 2004. Eligible patients were 18 years or older with any cancer
diagnosis, regardless of stage, and were about to receive their first chemotherapy
treatment containing doxorubicin and antiemetic prophylaxis with a 5-HT-receptor
antagonist, ondansetron, granisetron, or dolasetron plus dexamethasone or the
equivalent dose of intravenous methylprednisolone on the day of treatment.

7.2.2 Design

This study examined the relationship between expectancies and post-
chemotherapy nausea in a trial comparing antiemetic regimens for the treatment of
delayed nausea. First time chemotherapy patients completed a questionnaire about
their expectancies for nausea before their first infusion and then recorded their
nausea over four consecutive days following their treatment. As part of the trial,
patients were randomised to one of three antiemetic regimens for the second and
third days following their first chemotherapy infusion: Arm 1-prochlorperazine 10mg
p.o. every 8 hours, Arm 2-any first generation 5-HT₃ RA using standard dosage, or
Arm 3-prochlorperazine 10 mg p.o. as needed. However, only minimal differences
were observed among study arms (see Hickok et al., 2005 for additional details) and
this variable was controlled for as appropriate in the following analysis.

7.2.3 Measures

On-study Interview: During an interview conducted at the time of recruitment
patients were asked to provide information regarding demographics, diagnosis,
previous cancer-related treatment, and their history of nausea. The two questions
concerning history of nausea asked whether the patients were susceptible to motion
sickness and whether they experienced nausea during pregnancy. The latter question was scored as ‘Not applicable’ if the participant was male or had never been pregnant.

Expectancy Questionnaire: This questionnaire has previously been used by Roscoe et al. (2003) and contained questions assessing patient expectancies for posttreatment nausea, vomiting, fatigue, and hair loss. Four of these questions assessed expectancies for nausea. One required the patients to rate the likelihood that they would experience nausea after their chemotherapy treatment on a 5-point scale ranging from 1 (“I am certain I will not have this”) to 5 (“I am certain I will have this”). A second question required the patients to rate the expected severity of their post-chemotherapy nausea as “very mild or none at all”, “mild”, “moderate”, “severe”, “very severe”, or “intolerable”. A third question asked the patients to rate their perceived susceptibility to nausea compared with their friends and family as either “more”, “less”, or “the same”. A final question required patients to rate the likelihood of experiencing chemotherapy-related nausea compared with other cancer patients with the same diagnosis and undergoing the same treatment, again as “more”, “less”, or “the same”. These four expectancy questions were then combined by averaging z-scores to produce a single expectancy measure.

*Functional Assessment of Cancer Therapy Scale-General (FACT-G; version 4)*: This measure was used to assess patients’ quality of life. It contains 27-items assessing wellbeing across four domains: physical, functional, social/family, and emotional. The scores for each item are summed to produce an overall measure of global quality of life. The FACT-G is a widely used and well validated measure for
assessing quality of life in cancer patients undergoing chemotherapy treatment (Cella et al., 1993).

*Post-chemotherapy Nausea:* was assessed via a 4-day patient diary developed by Burish et al. (1987) and Carey and Burish (1988) specifically for this purpose. The diary required the patients to rate the severity of their nausea on a 7-point scale ranging from 1 (“not at all nauseated”) to 7 (“extremely nauseated”) for the morning, afternoon, evening, and night separately for each day. The diary was then used to calculate average nausea over the four days posttreatment as well as peak nausea, the highest severity rating for nausea at any time in the four days following treatment.

7.3.4 Procedure

For a full description of the procedures see Hickok et al. (2005). Briefly, after informed consent was gained, patients completed the on-study interview, the FACT-G, and the expectancy questionnaire. Patients then received their first chemotherapy infusion. Immediately following this they were given the 4-day diary to record their nausea. At the end of the fourth day the patients returned this diary and were given the FACT-G to complete for a second time. The study was approved by the institutional review board of the University of Rochester and every participating site approved the protocol in accordance with an assurance filed with and approved by the US Department of Health and Human Services.

7.3.5 Statistical Analyses

Two hierarchical regressions were used to determine the predictors of post-chemotherapy average and peak nausea in all patients. Age, gender, and susceptibility
to motion sickness comprised the first step, then diagnosis, followed by study arm, then pretreatment QoL, and finally expectancy. Diagnosis was dummy coded using breast cancer as the reference group and combining myeloma, endometrial, sarcoma, and bladder cancer patients into a single group because of their low numbers (see below). Nausea during pregnancy could not be included in this analysis because this question was only valid for previously pregnant women. In order to test whether this variable influenced the relationship between expectancy and nausea, the above analysis was repeated with the sub-sample of women who had been pregnant. Gender was excluded from this analysis as only women were involved. To assess the impact of level of expectancy on nausea four approximately equal groups were created using quartiles based on the combined expectancy measure. The groups were classified as not expectant (0-25<sup>th</sup> percentile), slightly expectant (26-50<sup>th</sup> percentile), somewhat expectant (50-75<sup>th</sup> percentile), and highly expectant (76-100<sup>th</sup> percentile). ANCOVA was then used to compare these groups with follow-up pairwise comparisons using Fisher’s LSD procedure.

A simultaneous regression was used to assess predictors of expectancy. This included age, gender, diagnosis, pretreatment quality of life, and susceptibility to motions sickness as possible predictors. As above, both of these analyses were repeated for the sub-sample of the previously pregnant women so as to determine whether the inclusion of nausea during pregnancy influenced the results. Finally, a hierarchical regression used to assess the impact of average and peak post-chemotherapy nausea on quality of life. Here, age and gender were entered as the first step, followed by diagnosis and study arm, then pretreatment quality of life, and then average and peak nausea in the final step. All statistical analyses were conducted
using SPSS software (version 15; SPSS Inc, Chicago, Ill) and results were considered significant when $p<.05$.

7.3 Results

7.3.1 Sample Characteristics

Six hundred and ninety-one patients enrolled in the study, of whom 671 provided evaluable data. Participants had an average age of 53 (range 25-90). The majority were female (94%), white (88%), and had received some college education (59%). Ninety percent had breast cancer, 9% had lymphoma, and the remaining 1% was a mix of myeloma, endometrial, sarcoma, and bladder cancer patients.

7.3.2 Overview of Nausea

Five hundred and sixty-two (84%) patients reported at least some nausea in the four days following treatment and 165 (25%) reported severe nausea (rating of 6 or 7 on the 7-point scale). Overall, average nausea over the four days following chemotherapy was 2.2 (SD=1.2) and the mean peak nausea was 4.0 (SD=2.1). Before their first chemotherapy treatment patients had a mean quality of life of 86.3 (SD=13.0), which decreased to 76.3 (SD=15.8) after the treatment.

7.3.3 Expectancy and Nausea

Table 7.1 shows the final step in the hierarchical regressions for average and peak nausea with all patients included. For average nausea, the overall model with age, gender, diagnosis, susceptibility to motion sickness, pretreatment quality of life, and expectancy included was significant and accounted for 16.7% of the variance, $R^2=.17$, $F(8,653)=14.6$, $p<.001$. Expectancy had a significant impact on average nausea. Specifically, an increase of one standard deviation on the expectancy measure
was associated with a .27 increase in average nausea after controlling for all other variables in the model, \( b = .27, t(653) = 4.44, p < .001 \). This meant that expectancy

Table 7.1. Final step in the hierarchical regression of predictors of post-chemotherapy nausea for all participants. †Dummy coded with breast cancer as the reference group. ††Dummy coded with Study Arm 1 as the reference group *Significant at \( p < .05 \), **Significant at \( p < .01 \), ***Significant at \( p < .001 \).

<table>
<thead>
<tr>
<th></th>
<th>( b )</th>
<th>SE</th>
<th>( \beta )</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A) Average Nausea</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age***</td>
<td>-.029</td>
<td>.004</td>
<td>-.273</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Gender</td>
<td>-.019</td>
<td>.252</td>
<td>-.004</td>
<td>.94</td>
</tr>
<tr>
<td>Motion Sickness</td>
<td>.033</td>
<td>.095</td>
<td>.013</td>
<td>.73</td>
</tr>
<tr>
<td>Lymphoma†</td>
<td>-.478</td>
<td>.201</td>
<td>-.117</td>
<td>.02</td>
</tr>
<tr>
<td>Other Diagnosis†</td>
<td>.571</td>
<td>.637</td>
<td>.032</td>
<td>.37</td>
</tr>
<tr>
<td>Study Arm 2†</td>
<td>.161</td>
<td>.003</td>
<td>.064</td>
<td>.12</td>
</tr>
<tr>
<td>Study Arm 3†</td>
<td>.046</td>
<td>.104</td>
<td>.018</td>
<td>.66</td>
</tr>
<tr>
<td>Pretreatment QoL*</td>
<td>-.008</td>
<td>.103</td>
<td>-.086</td>
<td>.03</td>
</tr>
<tr>
<td>Expectancy***</td>
<td>.271</td>
<td>.061</td>
<td>.176</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>B) Peak Nausea</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age***</td>
<td>-.050</td>
<td>.007</td>
<td>-.262</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Gender</td>
<td>-.280</td>
<td>.459</td>
<td>-.030</td>
<td>.54</td>
</tr>
<tr>
<td>Motion Sickness</td>
<td>.100</td>
<td>.173</td>
<td>.022</td>
<td>.57</td>
</tr>
<tr>
<td>Lymphoma†***</td>
<td>-.950</td>
<td>.366</td>
<td>-.129</td>
<td>.01</td>
</tr>
<tr>
<td>Other Diagnosis†</td>
<td>.782</td>
<td>1.16</td>
<td>.025</td>
<td>.50</td>
</tr>
<tr>
<td>Study Arm 2†</td>
<td>.173</td>
<td>.006</td>
<td>.038</td>
<td>.36</td>
</tr>
<tr>
<td>Study Arm 3†</td>
<td>.236</td>
<td>.189</td>
<td>.052</td>
<td>.21</td>
</tr>
<tr>
<td>Pretreatment QoL*</td>
<td>-.013</td>
<td>.189</td>
<td>-.076</td>
<td>.048</td>
</tr>
<tr>
<td>Expectancy***</td>
<td>.431</td>
<td>.111</td>
<td>.155</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

uniquely accounted for 2.5% of the variance in average nausea. Age, pretreatment quality of life, and having lymphoma were also significant predictors. Specifically, an increase in age of 10yrs corresponded to a decrease of 0.3 points on average nausea, \( b = -.03, t(653) = 7.32, p < .001 \). A 10-point increase in pretreatment quality of life corresponded to a .1 decrease in average nausea, \( b = -.008, t(653) = 2.25, p = .03 \).
Lymphoma patients reported average nausea that was .48 points lower than breast cancer patients, $b=-.478$, $t(653)=2.38$, $p=.02$.

When this analysis was repeated in the sub-sample of women who had been pregnant, expectancy remained a significant predictor, although it uniquely accounted for slightly less variance in average nausea, $R^2$ change=.019, $b=.25$, $t(537)=3.65$, $p<.001$. Nausea during pregnancy itself was significant, with those having experienced nausea during pregnancy reporting average nausea as .33 points higher than those who did not experience nausea during pregnancy, $b=.33$, $t(537)=3.35$, $p=.001$. The one difference in this analysis was that pretreatment quality of life was no longer significantly associated with average nausea, $t(537)=1.51$, $p>.05$.

For peak nausea, the overall model was also significant and accounted for 14.8% of the variance $F(8,653)=12.6$, $p<.001$. Expectancy was again significant after controlling for all other variables. Here, an increase of one standard deviation on the expectancy measure corresponded to a .43 increase in peak nausea, $b=-.43$, $t(653)=3.87$, $p<.001$. In this instance, expectancy uniquely accounted for 2.0% of the variance in peak nausea. As with average nausea, age, pretreatment quality of life, and having lymphoma were also significant predictors. A 10yr increase in age was associated with a .50 decrease in peak nausea, $b=-.05$, $t(653)=6.95$, $p<.001$. A 10-point increase in pretreatment quality of life corresponded to a .13 decrease in peak nausea, $b=-.013$, $t(653)=1.98$, $p<.048$. Lymphoma patients reported .95 points less peak nausea than breast cancer patients, $b=-.95$, $t(653)=2.60$, $p=.01$. 
As above, when this analysis was repeated on the sub-sample of previously pregnant women, expectancy remained a significant predictor and uniquely accounted for slightly less variance in peak nausea than when nausea during pregnancy was not included, $R^2$ change=.018, $b=.42, t(537)=3.41, p=.001$. Nausea during pregnancy was also significant. Women who experienced nausea during pregnancy had peak nausea .41 points higher than those not experiencing nausea during pregnancy, $b=.41, t(537)=2.28, p=.02$. Again, pretreatment quality of life failed to reach significance in this analysis, $t(537)=1.18, p>.05$.

### 7.3.4 Level of Expectancy and Nausea

Figure 7.1A shows average nausea across the different levels of expectancy. The ANCOVA, with age, gender, susceptibility to motion sickness, diagnosis, pretreatment quality of life, and study group as covariates, revealed that average nausea differed significantly as a function of level of expectancy, $F(3,653)=8.9, p<.001$. Pairwise comparisons using Fisher’s LSD revealed that, highly expectant patients reported significantly higher levels of average nausea than all other levels of expectancy. Specifically, highly expectant patients reported average nausea as .59 points higher than not expectant patients, $t(653)=4.51, p<.001$, .58 points higher than slightly expectant patients, $t(653)=4.54, p<.001$, and .47 points higher than somewhat expectant patients, $t(653)=3.78, p<.001$. There were no significant differences in average nausea between somewhat expectant, slightly expectant, and not expectant patients, highest $t(653)=1.02, p=.31$. The pattern of results was identical when the analysis when nausea during pregnancy was controlled for in the sub-sample of previously pregnant women.
Figure 7.1. Covariate adjusted mean (+SE) average (A) and peak (B) nausea by level of expectancy for all participants. Highly expectant individuals reported both more average nausea and higher peak nausea than all other expectancy levels, **p<.01, ***p<.001. No other differences were significant.
Figure 7.1B shows peak nausea for each level of expectancy. As with average nausea, the ANCOVA revealed significant differences in peak nausea across the different levels of expectancy after controlling for age, gender, motion sickness, study arm, diagnosis, and pretreatment quality of life, $F(3,653)=6.93$, $p<.001$. Highly expectant patients reported peak nausea .94 points higher than not expectant patients, $t(653)=3.94$, $p<.001$, .96 points higher than slightly expectant patients, $t(653)=4.13$, $p<.001$, and .65 higher than somewhat expectant patients, $t(653)=2.90$, $p=.004$. Again there were no differences between somewhat expectant, slightly expectant, and not expectant patients in terms of peak nausea, highest $t(653)=1.41$, $p=.16$. Also as with average nausea, this pattern of results was identical in the sub-sample of previously pregnant women.

7.3.5 Predictors of Expectancy

Table 7.2 displays the simultaneous regression predicting expectancies for nausea based on all participants. The overall model including age, gender, motion sickness, diagnosis, and pretreatment quality of life was significant and accounted for 18.6% of the variance in expectancy, $F(6,656)=24.9$, $p<.001$. The only two significant predictors were susceptibility to motion sickness and pretreatment quality of life. Controlling for all other variables, patients who reported being susceptible to nausea were, on average, .45 standard deviations higher on the expectancy measure than patients who reported not being susceptible to motion sickness, $b=.45$, $t(656)=7.75$, $p<.001$. An increase of 10 points in pretreatment quality of life predicted a decrease of .17 standard deviations on the expectancy measure, $b=.17$, $t(656)=7.75$, $p<.001$. This pattern did not change when the analysis was repeated for women who had been pregnant. However, nausea during pregnancy was also a significant predictor of
expectancy, such that women who had experienced nausea during pregnancy had expectancies for nausea that were .32 standard deviations higher, on average, than women who did not experience nausea during pregnancy, $b=.32$, $t(540)=5.12$, $p<.001$.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>$b$</th>
<th>SE</th>
<th>$\beta$</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.004</td>
<td>.003</td>
<td>-.060</td>
<td>.10</td>
</tr>
<tr>
<td>Gender</td>
<td>.239</td>
<td>.161</td>
<td>.072</td>
<td>.14</td>
</tr>
<tr>
<td>Motion Sickness***</td>
<td>.451</td>
<td>.058</td>
<td>.276</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lymphoma†</td>
<td>.195</td>
<td>.128</td>
<td>.074</td>
<td>.13</td>
</tr>
<tr>
<td>Other Diagnosis†</td>
<td>.570</td>
<td>.407</td>
<td>.050</td>
<td>.16</td>
</tr>
<tr>
<td>Pretreatment QoL***</td>
<td>-.017</td>
<td>.002</td>
<td>-.279</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

7.3.6 Nausea and Quality of Life

The overall model, with age, gender, pretreatment quality of life, average nausea, and peak nausea included was significant and accounted for 55% of the variability in post-chemotherapy quality of life, $R^2=.55$, $F(8,657)=89.4$, $p<.001$. Average nausea, peak nausea, and pretreatment quality of life were all significant predictors of posttreatment quality of life. This means that, even after controlling for age, gender, diagnosis, pretreatment quality of life, study group, and peak nausea, a 1-point increase in average nausea corresponded to a 5.2 point decrease in post-chemotherapy quality of life, $b=-5.1$, $t(656)=8.4$, $p<.001$. Similarly, a 1-point increase in peak nausea was associated with a .76 decrease in post-chemotherapy quality of life, $b=-.76$, $t(656)=2.22$, $p=.03$. A 1-point increase in pretreatment quality of life predicted a .58 increase in posttreatment quality of life, $b=.58$, $t(656)=17.6$, $p<.001$. Together, average and peak nausea uniquely accounted for 19.4% of the variability in post-chemotherapy quality of life, $R^2$ change=.19, $F(2,656)=142$, $p<.001$. Comparing
the standardised regression coefficients for average and peak nausea it was clear that average nausea had a much larger effect on post-chemotherapy nausea, $\beta = -0.389$ and $\beta = -0.102$ respectively.

### 7.4 Discussion

In this study, stronger pretreatment expectancies for nausea were significantly associated with higher average and peak nausea posttreatment. Importantly, this result was obtained after controlling for age, gender, susceptibility to motion sickness, and pretreatment quality of life in all patients and was maintained in the sub-sample in whom nausea during pregnancy could be included as a covariate. As such, these results show an independent contribution of expectancy to post-chemotherapy nausea and, thereby, provide strong support for expectancy as a causal factor in chemotherapy-related nausea. Although the current findings are consistent with previous studies that have found a relationship between expectancy and post-chemotherapy nausea (Haut et al., 1991; Jacobsen et al., 1988; Montgomery & Bovbjerg, 2000; Olver et al., 2005; Rhodes et al., 1995; Roscoe et al., 2004; Roscoe et al., 2000; Shelke et al., 2008), such studies have generally failed to adequately control for other potentially confounding variables and/or have used statistical techniques likely to overestimate the strength of this relationship. Thus, the current study provides the most convincing evidence to date that expectancies contribute to post-chemotherapy nausea.

A novel and interesting finding was that patients classified as highly expectant of nausea experienced higher levels of both average and peak nausea than all other levels of expectancy and that none of the other levels, somewhat expectant, slightly expectant, and not expectant, differed significantly from one another. This suggests
that patients who are highly expectant of experiencing nausea are particularly at risk of actually experiencing nausea after their chemotherapy treatment, whereas those with lower levels of expectancy share a similar decreased risk. This result was also found after controlling for age, gender, susceptibility to motion sickness, and quality of life, and again points to a causal influence of expectancy on post-chemotherapy nausea.

Another interesting finding was that susceptibility to motion sickness and nausea during pregnancy predicted expectancies for nausea. Although this is perhaps intuitive, it contradicts Montgomery and Bovbjerg’s (2003) finding that history of nausea was unrelated to expectancies for post-chemotherapy nausea. However, their analysis only involved a small sample for such an analysis, i.e. 31 patients, which may explain why they failed to find this effect. The current findings, therefore, suggest that patients’ experiences with nausea in other areas are likely to contribute to the nausea they expect to experience in response to chemotherapy. In addition to this, lower quality of life before treatment was associated with stronger expectancies for nausea. Here, the direction of causality is uncertain. It is possible that having strong expectancies for nausea following chemotherapy could detract from quality of life before treatment by creating apprehension and stress regarding the treatment. Conversely, it could be that those with lower quality of life before treatment are in a poorer state of health and therefore expect to be affected more by the chemotherapy treatment.

Increases in average nausea resulted in poorer quality of life posttreatment, a result consistent with previous research (Ballatori & Roila, 2003; Ballatori et al.,
Here, a 1-point increase in average nausea corresponded to a 5-point decrease in quality of life. Peak nausea also had a significant impact on quality of life after controlling for average nausea, although this effect was much smaller. This suggests that the continued presence of nausea has a more debilitating impact on chemotherapy patients’ quality of life compared with a bout of severe nausea. Average nausea and peak nausea independently of other variables accounted for 19.4% of the variability in post-chemotherapy QoL, a very large effect. Clearly, not only does nausea remain prevalent, it also continues to be a significant burden to cancer patients’ quality of life.

Given the current findings, patient expectancies reflect a possible point of intervention to reduce chemotherapy-induced nausea, as well as to improve cancer patients’ quality of life. A variety of studies in areas other than oncology have already shown that expectancy manipulations can have a beneficial impact on health outcomes. Perhaps most relevant is evidence that enhancing positive expectancies can decrease postoperative nausea in patients undergoing major gynaecological surgery (Williams et al., 1994) and protect against seasickness in naval cadets (Eden & Zuk, 1995). However, these types of manipulations may be more difficult to implement for cancer patients. First, it would be unethical to provide patients with unrealistically low expectancies regarding the likelihood of experiencing nausea as a result of their chemotherapy. Second, in addition to what cancer patients are told to expect from their treating health professionals, they will also draw on information from their family, friends, and the media. As such, any successful expectancy manipulation
would have to consider the source of the patient’s expectancies and the significance attributed to that source before attempting to adjust any maladaptive expectancies.

The finding that highly expectant patients experience more nausea than not expectant, slightly expectant, and somewhat expectant patients, and that these latter three did not differ significantly from each other may provide one answer. It suggests that completely eradicating expectancies for nausea is not necessary to gain a significant clinical improvement. For example, if a patient was initially highly expectant, but through discussion or some other manipulation became somewhat expectant, a significant decrease in both average nausea and peak nausea would be expected with a corresponding improvement in quality of life. Further reducing these expectancies to the level of slightly expectant, however, would produce only minimal additional benefit. Thus, a successful intervention might focus on patients who are highly expectant of experiencing nausea with the aim of reducing these maladaptive expectancies rather than removing all expectancy. In doing so, the patient would remain aware that there is a possibility of experiencing nausea, thereby avoiding any ethical concerns regarding giving patient’s unrealistically low expectations of experiencing nausea. Further, a small reduction in expectancy is likely to be much easier to achieve than convincing a highly expectant patient that he/she is very unlikely to experience nausea following chemotherapy.

One possibility along these lines is to include a discussion of the patient’s expectancies regarding side effects as part of the chemotherapy education. This would allow for the identification of highly expectant patients and provide an excellent opportunity to discuss and challenge these potentially maladaptive expectancies.
Given the relationship between expectancy and pretreatment quality of life, reducing strong expectancies for nausea during chemotherapy education may also improve patients’ pretreatment quality of life. To the best of my knowledge, patient expectancies are rarely, if ever, addressed in chemotherapy education.

These findings also have more general implications to do with placebo-induced side effects. Firstly, the present findings suggesting that higher expectancies for nausea actually contribute to post-chemotherapy nausea supports the possibility of placebo-induced side effects. To date, evidence of such effects has been limited, as reviewed in Chapter 5. Secondly, the fact that this effect was relatively small, with expectancy only uniquely accounting for about 2% of the variance in post-chemotherapy nausea, suggests that placebo-induced side effects could be relatively weak effects. This being the case, detecting placebo-induced side effects in this and other areas will likely require large sample size and this may partially explain the failure to find significant placebo-induced side effects in the previous study (see Chapter 6). Together with evidence of positive placebo effects, this supports the possibility raised in the previous chapter that providing informed consent or patient education is likely to lead to more benefit via the placebo effect than harm. However, some caution is required here as it did seem clear in the current study that strong expectancies for side effects can increase their likelihood in at least some patients. Thus, clinicians should remain mindful of how they present information about side effects to patients, as well as considering the extent to which patients already expect these side effects.
It is worth noting that, although susceptibility to motion sickness and nausea during pregnancy were included in the analysis, it is possible that the patients had other previous experiences of nausea, such as, medication-induced nausea. These other experiences of nausea may have affected patients’ expectancies for and/or actual experience of post-chemotherapy nausea, whether as a result of a general increased susceptibility to nausea or classical conditioning. If this were the case, then expectancy may simply have acted as a marker for other experiences with nausea rather than having a direct causal impact on post-chemotherapy nausea. However, the fact that expectancy was significantly associated with post-chemotherapy nausea after controlling for nausea during pregnancy, even though this predicted expectancy and both average and peak nausea itself, might suggest that this possibility is unlikely. A second potential limitation to the current findings is that nausea was assessed via self report. Although the diary used here has been used widely in this area (e.g. Burish et al., 1987; Carey & Burish, 1988; Roscoe et al., 2004; Shelke et al., 2008), it remains possible that this introduced some bias in patients’ reports of their nausea. As discussed in Chapter 5, this is a difficult problem to overcome when assessing placebo-induced side effects because many side effects are inherently subjective in nature. A third possible limitation is that the current study did not control for psychological factors, such as anxiety and emotional distress.

Overall, the current findings provide strong evidence that expectancy contributes to the development of post-chemotherapy nausea, which, in addition to being inherently unpleasant, also detracts from patients’ quality of life. Given that highly expectant patients appear at particular risk of experiencing post-chemotherapy nausea, an effective preventive strategy might be to target these patients. Including a
discussion of expectancies for nausea during chemotherapy education reflects a possibility that would have the added benefit of allowing direct exploration of the sources of these expectancies.
Chapter 8 – General Discussion

8.1 Summary

The current project assessed the placebo effect in double-blind RCTs and its contribution to treatment side effects, both of which have received relatively little research attention to date. The first study (Chapter 3) involved reanalysing a double-blind RCT of naltrexone and acamprosate for alcohol dependence and demonstrated that perceived treatment is a strong predictor of treatment outcomes in these types of trials. The role of perceived treatment in double-blind RCTs was explored further in Study 2 (Chapter 4) by developing an experimental model to test these effects. This involved two dummy (placebo) double-blind RCTs for cognitive performance. In these experiments, the bogus feedback given to participants about their performance heavily affected their perceived treatment, which in turn influenced their actual performance.

The two remaining studies investigated whether the placebo effect contributes to treatment side effects, labelled placebo-induced side effects. In the first of these (Chapter 6) a series of dummy trials for sleep difficulty were conducted, which involved manipulating the information participants received about potential, but bogus, side effects. Despite robust evidence for a placebo effect for sleep difficulty, these experiments produced minimal evidence for placebo-induced side effects. This raised the possibility that the placebo-induced side effects may exist, but that they are only small effects. The final study (Chapter 7) appeared to confirm this by showing that pretreatment expectancies for nausea predicted average and peak post-
chemotherapy nausea in a large sample of first time chemotherapy patients. The main implications of these findings are discussed below.

8.1.1 Perceived treatment in double-blind RCTs

The studies on perceived treatment in double-blind RCTs indicate a possible interplay between perceived treatment and observable outcomes. The development and application of an experimental model, in particular, demonstrated that perceived treatment can influence actual treatment outcomes via the placebo effect. If blinding fails, in that participants can guess their treatment allocation at rates better than chance, then the trial cannot validly determine whether the active treatment or the placebo effect led to any observed differences between study arms. Despite increasing acknowledgement of this problem (e.g. Benedetti, 2005; Benedetti, 2007; Day & Altman, 2000; Price, Finniss, & Benedetti, 2008), many researchers conducting double-blind RCTs fail to either test or report whether blinding was maintained (Fergusson et al., 2004; Hróbjartsson et al., 2007; Karanicolas et al., 2008). This means that the validity of these trials cannot be determined.

It is perhaps worth re-emphasising here that the aim of blinding is to evenly distribute expectancies across treatment groups and that this process is identical to randomising participants to study arms in an attempt to evenly distribute potentially confounding characteristics, such as, age, gender, and baseline severity of the condition. Seen in this light, the low rates of assessing blinding are highly peculiar, given that researchers almost always describe, compare, and control for differences in baseline characteristics across study arms that may influence treatment outcomes. Clearly researchers conducting double-blind RCTs should measure, assess, and
control for perceived treatment with the same consistency that they do for other potentially confounding patient characteristics. This would improve their ability to assess the validity of the trial’s outcome.

8.1.2 Information about treatment and its side effects

The experimental studies on placebo-induced side effects revealed that informing participants about side effects had little impact on their subsequent reports of adverse effects. This was despite a substantial proportion (around 30%) of participants who received placebo treatment freely reporting at least one side effect. The warning itself only had an effect on adverse effects in the third experiment, where, rather unexpectedly, both the treatment and no treatment groups reported changes in appetite that were consistent with the warning they had received. There was, however, a clear placebo effect for sleep difficulty, evidenced by the treatment group’s consistent reduced global sleep difficulty and increased total sleep time. This suggests that patients who are provided with information about the goal of a treatment and are warned about its possible side effects are likely to experience more benefit from this information than harm, via the placebo effect. If so, then the process of informed consent should add to the efficacy of a treatment while producing minimal placebo-induced side effects, as long as it contains information about the goal of the treatment. Nonetheless, the study involving first time chemotherapy patients provided evidence that expectancies for side effects can increase both their average and peak severity and that these adverse symptoms detracted from the patients’ quality of life. Thus, clinicians should be mindful of how information about side effects is presented and whether their patients enter treatment with strong or inaccurate preconceptions about its likely outcomes.
8.1.3 Importance of controlling for natural history

The studies assessing placebo-induced side effects in sleep difficulty also highlighted the importance of controlling for natural history when examining the placebo effect. In these studies, participants allocated to a no treatment control group reported reduced global sleep difficulty and shorter sleep latency. Many researchers have stressed the importance of controlling for natural history (e.g. Benedetti, 2009; Colloca & Benedetti, 2005; Kienle & Kiene, 1997; Kirsch & Sapirstein, 1998; Miller & Rosenstein, 2006; Price et al., 2008). While the vast majority of recent experimental studies do so, many non-experimental studies do not. This is particularly so for meta-analyses of changes in placebo groups in double-blind RCTs (e.g. Bittencourt et al., 2008; Fulda & Wetter, 2008; Macedo, Banos, & Farre, 2008; McCall et al., 2003). These studies typically involve calculating pooled effect sizes for improvement in patients receiving placebo from baseline to one or more follow-ups and erroneously claim that this reflects the placebo effect. However, it is clear from the three experiments on sleep difficulty that some participants receiving no treatment will report improvement and that an appropriate no treatment control group is required to differentiate changes that result from the placebo effect from those that result from the natural history of the condition.

8.1.4 Level of expectancy and the placebo effect

The finding that first time chemotherapy patients with a high level of expectancy (top quartile) for post-chemotherapy nausea experienced more nausea than those with all other levels of expectancy, but that there were no differences in nausea between the lower levels was very interesting. If this difference is attributable to expectancy, then it suggests that the magnitude of the placebo effect may not
necessarily increase linearly with the strength of the expectancy. Instead, it might be
the case that a certain level of expectancy is required to produce a placebo effect. If
so, this has important implications for developing interventions aimed at either
enhancing beneficial placebo effects or reducing adverse placebo effects. In the
former case, any expectancy manipulation may have to be powerful enough to create
sufficiently strong expectancies to cause a placebo effect. This could present a
significant challenge to these types of interventions. On the other hand, reducing
adverse placebo effects might require only a small reduction in expectancies, from
high to moderate, which is likely to be much more attainable than completely
eradicating expectancies for adverse outcomes. As a result, interventions aimed at
reducing adverse placebo effects might be more successful than those aimed at
enhancing beneficial placebo effects.

8.2 Limitations

Specific limitations do to with each study have been discussed in the relevant
chapters. However, there are two broader limitations that apply to these studies and to
research on the placebo effect more generally. The first concerns the extent to which
conveying uncertainty might undermine the placebo effect and the problems with
conveying certainty, namely ethicality and credibility. The second concerns the
relative advantages and disadvantages of measuring versus manipulating expectancies
when attempting to evaluate the placebo effect and how best to measure expectancy.

8.2.1 Uncertainty and the placebo effect

In order for an informational intervention to produce a placebo effect it must
elicit sufficiently strong expectancies for that effect. Conveying uncertainty may
undermine this. Thomas (1987) found that general practice patients who were given
placebo treatment and told that it was a medication certain to improve their symptoms reported significantly better outcomes than patients given the same placebo treatment but told that it was a medication that may or may not help them. This suggests that the magnitude of the placebo effect is sensitive to likelihood which is expressed about the efficacy of the treatment.

The side effect warning in the current experiments involving sleep difficulty suggested that participants may experience mild adverse symptoms. This uncertainty may have reduced the extent to which this information induced expectancies about these adverse effects and, thereby, might have undermined the placebo effect and could explain why they failed to produce any firm evidence for placebo-induced side effects. If so, then warnings about side effects that suggested a higher probability and of adverse effects may have led to stronger expectancies for these adverse effects and, therefore, increased likelihood of placebo-induced side effects.

However, there are questions regarding the credibility and ethicality of providing patients with information stating that a treatment is certain to be effective. Patients are likely to appraise the information they are provided with and whether it is consistent with their prior experience and any other information they have received. As such, a large proportion of patients are unlikely to find information declaring a treatment to work with absolute certainty credible. More importantly, although it may be acceptable in studies on healthy participants, providing patients with false information about the efficacy of their treatment would be unethical, especially if it undermined their ability to provide informed consent. This means that informational
manipulations will almost always contain some level of uncertainty and this may inhibit their ability to produce strong placebo effects.

8.2.2 Measuring versus manipulating expectancy

The current project used two approaches to examine the placebo effect. In the two studies on clinical populations, expectancies were measured and then their relationship with treatment outcomes was assessed. In the two studies on healthy volunteers, expectancy was experimentally manipulated via the information provided to these participants. This difference in approaches is also reflected in the placebo literature with some measuring expectancies (e.g. Andrykowski & Gregg, 1992; Linde et al., 2007; Olver et al., 2005; Roscoe et al., 2004; So, 2002) and others attempting to manipulate them (e.g. Amanzio & Benedetti, 1999; Benedetti, Amanzio et al., 1999; Levine, Stern, & Koch, 2006; Wolf, 1950).

Studies that involve manipulating expectancies are better able to determine whether the information has a causal impact on the placebo effect because of their experimental nature and might be considered superior for this reason. However, these studies rarely directly assess expectancies and, as such, are entirely reliant on the ability of the informational manipulation to influence expectancies. This leads to problems determining whether an unsuccessful informational manipulation failed because it did not induce expectancies or because there was no placebo effect, as occurred in the experimental studies on placebo-induced side effects. In the latter case, this would mean that the informational manipulation induced expectancies but that these expectancies did not affect outcomes. Studies that simply measure expectancies can directly assess the relationship between expectancy and outcomes, thereby overcoming problems to do with relying on informational manipulation.
affecting expectancies. However, these types of studies might be considered a weaker source of evidence for the placebo effect because they are correlational in nature.

A simple way to overcome this might be to include measures of expectancy in experiments with informational manipulations, which some studies have done (e.g. Fillmore & Vogel-Sprott, 1992; Kirsch & Weixel, 1988; Shelke et al., 2008). However, this is probably more complicated than it may seem. As mentioned in Chapter 6, questioning participants about expectancies during an experiment could influence their expectancies and/or make them question the true purpose of the study. Determining the best time to assess expectancies is also difficult. Measuring them immediately following the intervention and before the outcome is assessed will provide a prospective measure, but expectancies may change between these two time points especially if these are separated by a week or more. On the other hand assessing them immediately before or immediately after the outcomes are assessed could lead to priming that artificially inflates the strength of the relationship between expectancy and the outcome. In addition to this, no attempts have been made to determine the best method of measuring expectancies, with researchers often using Likert-type scales, but with different numbers of response options and different response categories (e.g. Fillmore & Vogel-Sprott, 1992; Jacobsen et al., 1988; Kirsch & Weixel, 1988; Roscoe et al., 2004), and at least one used 100mm visual analogue scales (Olver et al., 2005). This was no different in the current project with each study employing differing methods of assessing expectancies.

8.3 Future Directions

The implications and limitations to the current project suggest a number of interesting and important directions for future research. In terms of expectancies in
double-blind RCTs the experimental model developed here could serve as a useful method of further exploring the relationship between perceived treatment and outcomes. Of interest would be whether the same results would be obtained if the observable change manipulated involved a side effect. Such a study could determine whether side effects do influence perceived treatment, beyond the established correlational evidence (see Shapiro & Shapiro, 1997b for a review). If so, then this would provide a strong argument for incorporating active placebos into double-blind RCTs and may influence ethical considerations regarding the benefits and costs of doing so.

The experimental study on placebo-induced side effects highlighted potential difficulties associated with employing informational manipulations. Future studies could investigate how the content of the information affects expectancies and the placebo effect. In fact it was my intention to use dummy trials for sleep difficulty to compare the magnitude of placebo-induced side effects when participants are provided with high (60%) versus low (30%) probabilities about side effects. Unfortunately the failure to detect clear placebo-induced side effects did not permit this. However, such a study could be conducted in an area where there is robust evidence for the placebo effect, such as pain. This would provide confirmation of Thomas’ (1987) finding regarding conveying uncertainty and might indicate how best to develop a successful informational manipulation. Of course, in doing so researchers should be mindful of the trade-off between conveying certainty and the associated problems to do with credibility and ethicality.
The study on first time chemotherapy patients’ expectancies and actual experience of nausea suggested that interventions targeted at patients with particularly high expectancies for nausea might be most successful. As mentioned earlier, one such intervention might be to incorporate a discussion of expectancies into the chemotherapy education session that most patients undergo. During this education session any overly strong expectancies for nausea could be challenged.

The issues to do with measuring versus manipulating expectancies could also be addressed empirically. In order to determine the possible influence of questioning participants about their expectancies, an experimental study involving an informational manipulation could have one group that is asked to rate their expectancies and another that is not. A similar design could also be used to determine whether the timing of the expectancy assessment affects outcomes. One such study could involve four groups. The first group would be asked to report their expectancies immediately following the informational manipulation. The second group would be asked immediately before reporting their treatment outcomes. The third group could be asked after reporting their treatment outcomes. The fourth group could be asked at all of these time points, which would enable examination of the effects of repeatedly questioning participants about their expectancies. Finally, it would be highly useful to attempt to compare actual measures of assessing expectancies, for example Likert-type scales versus visual analogue scales, and whether a single item is sufficient or whether multiple items should be incorporated.
8.4 Conclusions

The current findings provided firm evidence that the placebo effect influences treatment outcomes in double-blind RCTs and some evidence that expectancy can increase the occurrence and severity of treatment side effects. These findings highlight the importance of considering the possible influence of patient expectancies when delivering medical treatment. Specifically, researchers conducting double-blind RCTs should measure and, where necessary, control for participants’ perceived treatment in order to ensure that any differences between the active treatment and the placebo can be attributed to the active treatment alone. In terms of side effects, reducing any unrealistically high expectancies for adverse effects through discussion, or some other intervention, could reduce the burden that these effects have on patients, which could in turn improve compliance and treatment outcomes.
References


hypertension: observations from a Department of Veterans Affairs Cooperative Study. *Archives of Internal Medicine, 160*, 1449-1454.


Appendices

Table of Contents

Appendix 1 - List of published and submitted peer-reviewed journal articles based on work undertaken for the degree of Doctor of Philosophy

Appendix 2 – List of conference presentations based on work undertaken for the degree of Doctor of Philosophy

Appendix 3 – Information sheets and study-specific questionnaires for Study 2: Perceived treatment, false feedback, and cognitive performance: Two dummy double-blind RCTs

3a – Information Sheet for Study 2, Experiment 1
3b – Beliefs Questionnaire for Study 2, Experiment 1
3c – Information Sheet for Study 2, Experiment 2 (Double-blind Group)
3d – Information Sheet for Study 2, Experiment 2 (Control Group)
3e - Beliefs Questionnaire for Study 2, Experiment 2 (Double-blind Group)
3f - Beliefs Questionnaire for Study 2, Experiment 2 (Control Group)

Appendix 4 – Information sheets and study-specific questionnaires for Study 3: Information and Placebo-Induced Side Effects: Three Dummy Trials for Sleep Difficulty

4a – Information Sheet for Study 3, Experiment 1 (Suggestion)
4b – Posttreatment Questionnaire for Study 3, Experiment 1
4c – Adverse Symptom Checklist for Study 3, Experiment 1
4d – Information Sheet for Study 3, Experiment 2 (Suggestion A)
4e – Information Sheet for Study 3, Experiment 2 (Suggestion B)
4f – Posttreatment Questionnaire for Study 3, Experiment 2
4g – Adverse Symptom Checklist for Study 3, Experiment 2
4h – Information Sheet for Study 3, Experiment 3 (Increase)
4i – Information Sheet for Study 3, Experiment 3 (Decrease)
4j – Posttreatment Questionnaire for Study 3, Experiment 3
4k – Adverse Symptom Checklist for Study 3, Experiment 3
Appendix 1 - List of published and submitted peer-reviewed journal articles based on work undertaken for the degree of Doctor of Philosophy


Appendix 2 – List of conference presentations based on work undertaken for the degree of Doctor of Philosophy

**Experimental Psychology Conference**, Wollongong, April 2008 – “Perceived treatment allocation in double-blind randomised placebo-controlled trials”


**Sydney Cancer Conference**, Sydney, July 2008 – “Expectancy, post-chemotherapy nausea, and QoL”

**Sydney University Postgraduate Psychology Conference**, Sydney, October 2006 – “Expectancy as a mediating factor in double-blind RCTs”  
[ Awarded Best Presentation]
Appendix 3a – Information Sheet for Study 2, Experiment 1.

Prof. R A Boakes
School of Psychology
The University of Sydney

Caffeine Effects

Information for Participants

Introduction
You are invited to take part in a research study investigating the temporary effects of caffeine on cognitive performance. Caffeine is the world’s most consumed drug and may serve as a relatively cheap and safe short term cognitive enhancer. In this study, we are interested in whether caffeine will improve your performance on a cognitive task that requires sustained attention. The study is a double-blind placebo controlled trial of caffeine, so you will be allocated to receive either caffeine (150mg) or a placebo pill, but you will not be told which one you have been given and the researcher will also not know your allocation. To assess the effects of caffeine versus placebo you will be asked to complete a series of cognitive tasks over the course of the session.

This study is being conducted by:
Prof. Robert Boakes (Chief Investigator),
University of Sydney
Mr Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to undergo the following procedures:

1) Give some personal details (e.g. age, sex) and fill out a questionnaire about your usual caffeine intake habits.
2) Complete a computer based sustained attention task (3 x 5min).
3) Complete a questionnaire involving questions about your experiences while you were performing the cognitive task.

Precautions
You should not take part in this study if
1) you have diabetes
2) have high blood pressure
3) have a serious heart condition
4) are currently taking medication (excluding the contraceptive pill)
Benefits
You will be awarded one hour of credit for participating in this experiment which will contribute 1.25% to the total 5% credit allocated to experiment participation for Psychology 1002.

Remember
Your participation in this study is entirely voluntary. **You may withdraw at anytime without having to give a reason.** Whatever your decision, please be assured that it will not have any repercussions.

When you have read this information, Ben Colagiuri will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact him on 0402 236 400, or by email at benc@psych.usyd.edu.au. This information sheet is for you to keep.

Any person with concerns or complaints about the conduct of this or any other study can contact the secretary of the Human Ethics Committee, University of Sydney, on (02) 9351 4811.
Appendix 3b - Beliefs Questionnaire for Study 2, Experiment 1.

Caffeine Study

Caffeine Effects Questionnaire

1) Do you think that you were given the caffeine pill or the placebo pill?

| CAFFEINE | PLACEBO |

2) On a scale from 0 – 10, how certain are you that this is the medication you are receiving?

| NOT AT ALL | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | EXTREMELY CERTAIN |

3) Are there any particular reasons why you believe this?

_____________________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

GO TO NEXT PAGE
4) Do you remember your accuracy scores for the three tests?

   Baseline  _______________

   Posttest 1 _______________

   Posttest 2 _______________

5) Please list any caffeinated products that you have consumed so far today?

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________
Appendix 3c – Information Sheet for Study 2, Experiment 2. (Double-blind Group)

Prof. R A Boakes
School of Psychology
The University of Sydney

Piracetam Trial

Information for Participants

Introduction
You are invited to take part in a research study investigating the temporary effects of piracetam on cognitive performance and heart rate. Piracetam is an over-the-counter nootropic or “smart drug”. It has been shown to increase learning and memory when used consistently over a period of time. However, there is also some evidence to suggest that a single dose of piracetam can produce immediate short term improvements in cognitive performance. The most likely mechanism for this improved cognitive performance is an increase in blood flow via increased heart rate which provides more oxygen to the brain. In this study, we are interested in whether a single dose of piracetam will improve your cognitive performance on a task involving working memory and sustained attention and increase your heart rate. The study is a double-blind placebo controlled trial of piracetam, so you will be allocated to receive either piracetam (150mg) or a placebo pill, but you will not be told which one you have been given and the researcher will also not know your allocation. To assess the immediate effects of piracetam versus placebo you will be asked to complete a cognitive task and have your heart rate measured three times over the course of this session.

This study is being conducted by:
Prof. Robert Boakes (Chief Investigator),
University of Sydney
Mr Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to undergo the following procedures:
1) Provide some personal details (e.g. age, sex).
2) Take one piracetam or placebo tablet, although you will not know which one you have been given.
3) Have your heart rate measured using a blood volume pulse sensor (3 times).
4) Complete a computer based sustained attention task (3x5min).
5) Complete a questionnaire involving questions about the effect of the treatment you received.
Benefits
You will be awarded one hour of credit for participating in this experiment which will contribute 1.25% to the total 5% credit allocated to experiment participation for Psychology 1002.

Remember
Your participation in this study is entirely voluntary. You may withdraw at anytime without having to give a reason. Whatever your decision, please be assured that it will not have any repercussions.

When you have read this information, Ben Colagiuri will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact him on 0402 236 400, or by email at benc@psych.usyd.edu.au. This information sheet is for you to keep.

Any person with concerns or complaints about the conduct of this or any other study can contact the secretary of the Human Ethics Committee, University of Sydney, on (02) 9351 4811.
Appendix 3d – Information Sheet for Study 2, Experiment 2. (Control Group)

Prof. R A Boakes
School of Psychology
The University of Sydney

Piracetam Trial
Information for Participants

Introduction
You are invited to take part in a research study investigating the temporary effects of piracetam on cognitive performance and heart rate. Piracetam is an over-the-counter nootropic or “smart drug”. It has been shown to increase learning and memory when used consistently over a period of time. However, there is also some evidence to suggest that a single dose of piracetam can produce immediate short term improvements in cognitive performance. The most likely mechanism for this improved cognitive performance is an increase in blood flow via increased heart rate which provides more oxygen to the brain. In this study, we are interested in whether a single dose of piracetam will improve your cognitive performance on a task involving working memory and sustained attention and increase your heart rate. You will be allocated either to receive piracetam (150mg) or to a no treatment control group that does not receive the piracetam. To assess the immediate effects of piracetam versus no treatment you will be asked to complete a cognitive task and have your heart rate measured three times over the course of this session.

This study is being conducted by:
Prof. Robert Boakes (Chief Investigator),
University of Sydney
Mr Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to undergo the following procedures:
1) Provide some personal details (e.g. age, sex).
2) Be allocated to either receive a piracetam pill or no treatment.
3) Have your heart rate measured using a blood volume pulse sensor (3 times).
4) Complete a computer based sustained attention task (3x5min).
5) Complete a questionnaire involving questions about the effect of the treatment you received.
Benefits
You will be awarded one hour of credit for participating in this experiment which will contribute 1.25% to the total 5% credit allocated to experiment participation for Psychology 1002.

Remember
Your participation in this study is entirely voluntary. **You may withdraw at anytime without having to give a reason.** Whatever your decision, please be assured that it will not have any repercussions.

When you have read this information, Ben Colagiuri will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact him on 0402 236 400, or by email at benc@psych.usyd.edu.au. This information sheet is for you to keep.

Any person with concerns or complaints about the conduct of this or any other study can contact the secretary of the Human Ethics Committee, University of Sydney, on (02) 9351 4811.
Appendix 3e - Beliefs Questionnaire for Study 2, Experiment 2. (Double-blind Group)

Professor R A Boakes  
School of Psychology  
The University of Sydney

---

**Piracetam Trial**

**Piracetam Effects Questionnaire**

1) Do you think that you were given the piracetam pill or the placebo pill?

<table>
<thead>
<tr>
<th>PIRACETAM</th>
<th>PLACEBO</th>
</tr>
</thead>
</table>

2) On a scale from 0 – 10, how certain are you that this was the medication that you received?

<table>
<thead>
<tr>
<th>NOT AT ALL</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>EXTREMELY CERTAIN</th>
</tr>
</thead>
</table>

3) Are there any particular reasons why you believe this?

________________________
________________________
________________________
________________________

GO TO NEXT PAGE
4) Do you remember your accuracy scores for the three tests?

Baseline ____________________

Posttest 1 ________________

Posttest 2 ________________

5) Do you have any additional comments about the study?

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

______________________________________________________________________

END OF QUESTIONNAIRE
Appendix 3f - Beliefs Questionnaire for Study 2, Experiment 2. (Control Group)

Professor R A Boakes
School of Psychology
The University of Sydney

Piracetam Trial

Piracetam Effects Questionnaire

1) Did you receive piracetam or no treatment?

PIRACETAM   NO TREATMENT

2) Do you remember your accuracy scores for the three tests?

Baseline   _______________
Posttest 1  _______________
Posttest 2  _______________

3) Do you have any additional comments about the study?

_____________________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

END OF QUESTIONNAIRE
Appendix 4a – Information Sheet for Study 3, Experiment 1 (Suggestion)

Sleep Trial
Information for Participants

Introduction
You are invited to take part in a research study investigating the effectiveness of a new herbal medication, SX3752, designed to reduce difficulty sleeping. Difficulty sleeping is one of the most common general health problems reported today. It includes taking a long time to get to sleep, waking up frequently, and waking up earlier than expected and not being able to return to sleep. The main purpose of this study is to assess the effectiveness of SX3752 for reducing difficulty sleeping over a one week period. SX3752 is a pill that you take by mouth with a glass of water each night before you go to bed. It is believed to be fast acting and to have the potential to reduce common sleep problems, such as, difficulty getting to sleep, waking continually throughout the night, and waking early in the morning. If you choose to participate in this study, you may be asked to take one SX3752 tablet each night for one week, or you may be assigned to a no-treatment control group, in which case you would not actually receive any SX3752 tablets.

This study is being conducted by:
Prof. Phyllis Butow (Chief Investigator),
University of Sydney
Prof. Robert Boakes,
University of Sydney
Mr. Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to:

Initial Session (Approximately 30min)
1) Provide some personal details, e.g. age, gender.
2) Complete a questionnaire about your sleeping habits over the past month.
3) Complete a questionnaire about some of your views on medical treatment.

Over the next week
1) Take one SX3752 tablet each night before you go to bed for the next week. You take these tablets by mouth with a glass of water.

At the end of the week (Take-home Questionnaires, Approximately 30min)
1) Complete a questionnaire about the quality of your sleep over the week.
2) Complete a questionnaire about how effective you think the SX3752 tablets are.
3) Complete a questionnaire about your general health symptoms over the week.
4) Return these questionnaires within one week of the final day of your treatment.
Risks
Some people who take SX3752 report some minor side-effects. These may include feeling more sleepy than usual when you awaken, having a dry mouth, and mild headaches. If you do experience these side effects, they will go away completely when you stop taking SX3752.

Benefits
While we intend that this research study furthers medical knowledge and may help people experiencing difficulty sleeping, it may not be of direct benefit to you.

Costs
Participation in this study will not cost you anything, nor will you be paid.

Voluntary Participation
Participation in this study is entirely voluntary. You do not have to take part in it. If you do take part, you can withdraw at any time without having to give a reason. Whatever your decision, please be assured that it will not have any repercussions.

Confidentiality
All the information collected from you for the study will be treated confidentially, and only the researchers named above will have access to it. The study results may be presented at a conference or in a scientific publication, but individual participants will not be identifiable in such a presentation.

Further Information
When you have read this information, Ben Colagiuri will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact him on 0402 236 400, or by email at benc@psych.usyd.edu.au.

This information sheet is for you to keep.

Any person with concerns or complaints about the conduct of a research study can contact the Senior Ethics Officer, Ethics Administration, University of Sydney, on (02) 9351 4811 (Telephone); (02) 9351 6706 (Facsimile) or gbrody@mail.usyd.edu.au (Email).
Appendix 4b – Posttreatment Questionnaire for Study 3, Experiment 1

Post Treatment Questionnaire

Instructions:
Please answer the following questions about your treatment with SX3752 over the last week as honestly as possible.

1. How effective were the SX3752 tablets for reducing your sleep difficulty?
   
   NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10 VERY EFFECTIVE

2. At the beginning of the study, how effective did you expect the SX3752 tablets would be for reducing your sleep difficulty?

   NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10 VERY EFFECTIVE

3. Did you take one SX3752 tablet every night?   YES    NO
   If no, how many nights did you miss taking the SX3752 tablet_______
   What led to you not taking the SX3752 tablets?

   ____________________________________________________________

   ____________________________________________________________

4. Did you experience any side effects while you were taking the SX3752 tablets?   YES    NO
   If yes, what were these side effects and how severe were they?

   Side effect 1 (please specify) __________________________________
   VERY MILD  0  1  2  3  4  5  6  7  8  9  10 VERY SEVERE

   Side effect 2 (please specify) __________________________________
   VERY MILD  0  1  2  3  4  5  6  7  8  9  10 VERY SEVERE

   Side effect 3 (please specify) __________________________________
   VERY MILD  0  1  2  3  4  5  6  7  8  9  10 VERY SEVERE
Appendix 4c – Adverse Symptom Checklist for Study 3, Experiment 1

General Health Questionnaire

Instructions:
Please answer the following questions about your general health over the past week as accurately as possible. Try to consider the entire week and not just the last couple of days.

1. During the last week, have you felt more sleepy than usual when you wake up? YES NO
   If yes, on how many days did this happen _________
   On average, how much more sleepy did you feel than usual?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 VERY MUCH MORE

2. During the last week, have you felt more tired than usual in the evenings? YES NO
   If yes, on how many days did this happen _________
   On average, how much more tired did you feel than usual?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 VERY MUCH MORE

3. During the last week, have you experienced any drowsiness? YES NO
   If yes, on how many days did you feel drowsy _________
   On average, how severe would you rate this drowsiness?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

4. During the last week, have you experienced any headaches? YES NO
   If yes, on how many days did you experience headaches _________
   On average, how severe would you rate these headaches?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

5. During the last week, have you experienced any dizziness? YES NO
   If yes, on how many days did you feel dizzy _________
   On average, how severe would you rate this dizziness?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE
6. During the last week, have you experienced any nausea?  YES  NO
   If yes, on how many days did you experience nausea _________
   On average, how severe would you rate the nausea?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

7. During the last week, have you experienced a dry mouth?  YES  NO
   If yes, on how many days did you have a dry mouth_________
   On average, how severe would you rate this dry mouth?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

8. During the last week, have you felt more restless than usual?  YES  NO
   If yes, on how many days did you feel more restless than usual _________
   On average, how much more restless did you feel than usual?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 MUCH MORE

9. During the last week, have you felt more irritable than usual?  YES  NO
   If yes, on how many days did you feel more irritable than usual _________
   On average, how much more irritable did you feel than usual?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 MUCH MORE

10. Did you experience any other general health symptoms during the past week?  YES  NO
   If yes, what were these symptoms ____________________________
   On how many days did you experience these symptoms _________
   On average, how severe would you rate these symptoms?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 VERY SEVERE

11. Do you have any other comments about your general health over the last week?

_________________________________________________________________
_________________________________________________________________
_________________________________________________________________

END OF QUESTIONNAIRE
Please check that you have answered ALL questions.
Appendix 4d – Information Sheet for Study 3, Experiment 2 (Suggestion A)

Sleep Trial
Information for Participants

Introduction
You are invited to take part in a research study investigating the effectiveness of a new herbal medication, SX3752, designed to reduce difficulty sleeping. Difficulty sleeping is one of the most common general health problems reported today. It includes taking a long time to get to sleep, waking up frequently, and waking up earlier than expected and not being able to return to sleep. The main purpose of this study is to assess the effectiveness of SX3752 for reducing difficulty sleeping over a one week period. SX3752 is a pill that you take by mouth with a glass of water each night before you go to bed. It is believed to be fast acting and to have the potential to reduce common sleep problems, such as, difficulty getting to sleep, waking continually throughout the night, and waking early in the morning. If you choose to participate in this study, you may be asked to take one SX3752 tablet each night for one week, or you may be assigned to a no-treatment control group, in which case you would not actually receive any SX3752 tablets.

This study is being conducted by:
Prof. Phyllis Butow (Chief Investigator),
University of Sydney
Prof. Robert Boakes,
University of Sydney
Mr. Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to:

Initial Session (Approximately 30min)
1) Provide some personal details, e.g. age, gender.
2) Complete a questionnaire about your sleeping habits over the past month.
3) Complete a questionnaire about some of your views on medical treatment.

Over the next week
1) Take one SX3752 tablet each night before you go to bed for the next week. You take these tablets by mouth with a glass of water.

At the end of the week (Take-home Questionnaires, Approximately 30min)
1) Complete a questionnaire about the quality of your sleep over the week.
2) Complete a questionnaire about how effective you think the SX3752 tablets are.
3) Complete a questionnaire about your general health symptoms over the week.
4) Return these questionnaires within one week of the final day of your treatment.
Risks
Some people who take SX3752 report some minor side-effects. These may include feeling drowsy, having a dry mouth, and nausea. If you do experience these side effects, they will go away completely when you stop taking SX3752.

Benefits
While we intend that this research study furthers medical knowledge and may help people experiencing difficulty sleeping, it may not be of direct benefit to you.

Costs
Participation in this study will not cost you anything, nor will you be paid.

Voluntary Participation
Participation in this study is entirely voluntary. You do not have to take part in it. If you do take part, you can withdraw at any time without having to give a reason. Whatever your decision, please be assured that it will not have any repercussions.

Confidentiality
All the information collected from you for the study will be treated confidentially, and only the researchers named above will have access to it. The study results may be presented at a conference or in a scientific publication, but individual participants will not be identifiable in such a presentation.

Further Information
When you have read this information, Ben Colagiuri will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact him on 0402 236 400, or by email at benc@psych.usyd.edu.au.

This information sheet is for you to keep.

Any person with concerns or complaints about the conduct of a research study can contact the Senior Ethics Officer, Ethics Administration, University of Sydney, on (02) 9351 4811 (Telephone); (02) 9351 6706 (Facsimile) or gbrody@mail.usyd.edu.au (Email).
Appendix 4e – Information Sheet for Study 3, Experiment 2 (Suggestion B)

Sleep Trial
Information for Participants

Introduction
You are invited to take part in a research study investigating the effectiveness of a new herbal medication, SX3752, designed to reduce difficulty sleeping. Difficulty sleeping is one of the most common general health problems reported today. It includes taking a long time to get to sleep, waking up frequently, and waking up earlier than expected and not being able to return to sleep. The main purpose of this study is to assess the effectiveness of SX3752 for reducing difficulty sleeping over a one week period. SX3752 is a pill that you take by mouth with a glass of water each night before you go to bed. It is believed to be fast acting and to have the potential to reduce common sleep problems, such as, difficulty getting to sleep, waking continually throughout the night, and waking early in the morning. If you choose to participate in this study, you may be asked to take one SX3752 tablet each night for one week, or you may be assigned to a no-treatment control group, in which case you would not actually receive any SX3752 tablets.

This study is being conducted by:
Prof. Phyllis Butow (Chief Investigator),
University of Sydney
Prof. Robert Boakes,
University of Sydney
Mr. Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to:

Initial Session (Approximately 30min)
1) Provide some personal details, e.g. age, gender.
2) Complete a questionnaire about your sleeping habits over the past month.
3) Complete a questionnaire about some of your views on medical treatment.

Over the next week
1) Take one SX3752 tablet each night before you go to bed for the next week. You take these tablets by mouth with a glass of water.

At the end of the week (Take-home Questionnaires, Approximately 30min)
1) Complete a questionnaire about the quality of your sleep over the week.
2) Complete a questionnaire about how effective you think the SX3752 tablets are.
3) Complete a questionnaire about your general health symptoms over the week.
4) Return these questionnaires within one week of the final day of your treatment.
Risks
Some people who take SX3752 report some minor side-effects. These may include feeling dizzy, having blurred vision, and getting sore eyes. If you do experience these side effects, they will go away completely when you stop taking SX3752.

Benefits
While we intend that this research study furthers medical knowledge and may help people experiencing difficulty sleeping, it may not be of direct benefit to you.

Costs
Participation in this study will not cost you anything, nor will you be paid.

Voluntary Participation
Participation in this study is entirely voluntary. You do not have to take part in it. If you do take part, **you can withdraw at any time without having to give a reason.** Whatever your decision, please be assured that it will not have any repercussions.

Confidentiality
All the information collected from you for the study will be treated confidentially, and only the researchers named above will have access to it. The study results may be presented at a conference or in a scientific publication, but individual participants will not be identifiable in such a presentation.

Further Information
When you have read this information, Ben Colagiuri will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact him on 0402 236 400, or by email at benc@psych.usyd.edu.au.

This information sheet is for you to keep.

---

Any person with concerns or complaints about the conduct of a research study can contact the Senior Ethics Officer, Ethics Administration, University of Sydney, on (02) 9351 4811 (Telephone); (02) 9351 6706 (Facsimile) or gbriody@mail.usyd.edu.au (Email).
Appendix 4f – Posttreatment Questionnaire for Study 3, Experiment 2

Post Treatment Questionnaire

Instructions:
Please answer the following questions about your treatment with SX3752 over the last week as honestly as possible.

1. How effective **were** the SX3752 tablets for reducing your sleep difficulty?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 VERY EFFECTIVE

2. At the beginning of the study, how effective did you **expect** the SX3752 tablets would be for reducing your sleep difficulty?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 VERY EFFECTIVE

3. Did you take one SX3752 tablet **every night**? YES  NO
   If no, how many nights did you miss taking the SX3752 tablet_________
   What led to you not taking the SX3752 tablet?
   ____________________________________________________

4. Did you experience any **side effects** while you were taking the SX3752 tablets?  YES  NO
   If yes, what were these side effects and how severe were they?
   Side effect 1 (please specify) __________________________________________
   VERY MILD 0 1 2 3 4 5 6 7 8 9 10 VERY SEVERE
   Side effect 2 (please specify) __________________________________________
   VERY MILD 0 1 2 3 4 5 6 7 8 9 10 VERY SEVERE
   Side effect 3 (please specify) __________________________________________
   VERY MILD 0 1 2 3 4 5 6 7 8 9 10 VERY SEVERE
General Health Questionnaire

Instructions:
Please answer the following questions about your general health over the past week as accurately as possible. Try to consider the entire week and not just the last couple of days.

1. During the last week, have you felt more sleepy than usual when you wake up?  YES     NO
   If yes, on how many days did this happen _________
   On average, how much more sleepy did you feel than usual?
   NOT AT ALL     0     1     2     3     4     5     6     7     8     9     10     VERY MUCH MORE

2. During the last week, have you felt more tired than usual in the evenings?  YES     NO
   If yes, on how many days did this happen _________
   On average, how much more tired did you feel than usual?
   NOT AT ALL     0     1     2     3     4     5     6     7     8     9     10     VERY MUCH MORE

3. During the last week, have you experienced any drowsiness?  YES     NO
   If yes, on how many days did you feel drowsy _________
   On average, how severe would you rate this drowsiness?
   NOT AT ALL     0     1     2     3     4     5     6     7     8     9     10     EXTREMELY SEVERE

4. During the last week, have you experienced any dizziness?  YES     NO
   If yes, on how many days did you feel dizzy _________
   On average, how severe would you rate this dizziness?
   NOT AT ALL     0     1     2     3     4     5     6     7     8     9     10     EXTREMELY SEVERE

5. During the last week, have you experienced any headaches?  YES     NO
   If yes, on how many days did you experience headaches _________
   On average, how severe would you rate these headaches?
   NOT AT ALL     0     1     2     3     4     5     6     7     8     9     10     EXTREMELY SEVERE
6. During the last week, have you experienced any blurred vision?  YES  NO
If yes, on how many days did you have blurred vision __________
On average, how severe would you rate this blurred vision?
NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  EXTREMELY SEVERE

7 During the last week, have you experienced sore eyes?  YES  NO
If yes, on how many days did you have sore eyes __________
On average, how severe would you rate these sore eyes?
NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  EXTREMELY SEVERE

8. During the last week, have you experienced a dry mouth?  YES  NO
If yes, on how many days did have a dry mouth_________
On average, how severe would you rate this dry mouth?
NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  EXTREMELY SEVERE

9. During the last week, have you experienced a sore throat?  YES  NO
If yes, on how many days did you experience sore throat __________
On average, how severe would you rate this sore throat?
NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  EXTREMELY SEVERE

10. During the last week, have you experienced any nausea?  YES  NO
If yes, on how many days did you experience nausea __________
On average, how severe would you rate the nausea?
NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  EXTREMELY SEVERE

11. During the last week, have you experienced any stomach cramps?  YES  NO
If yes, on how many days did you experience stomach cramps __________
On average, how severe would you rate these stomach cramps?
NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  EXTREMELY SEVERE
12. Did you experience any other general health symptoms during the past week?   YES   NO

   If yes, what were these symptoms ____________________________

   On how many days did you experience these symptoms _________

   On average, how severe would you rate these symptoms?
   NOT AT ALL   0      1      2      3      4      5      6      7      8      9      10    VERY SEVERE

13. Do you have any other comments about your general health over the last week?

___________________________________________________________________

___________________________________________________________________

___________________________________________________________________

END OF QUESTIONNAIRE
Please check that you have answered ALL questions.
Appendix 4h – Information Sheet for Study 3, Experiment 3 (Increase)

Sleep Trial
Information for Participants

Introduction
You are invited to take part in a research study investigating the effectiveness of a new herbal medication, SX3752, designed to reduce difficulty sleeping. Difficulty sleeping is one of the most common general health problems reported today. It includes taking a long time to get to sleep, waking up frequently, and waking up earlier than expected and not being able to return to sleep. The main purpose of this study is to assess the effectiveness of SX3752 for reducing difficulty sleeping over a one week period. SX3752 is a pill that you take by mouth with a glass of water each night before you go to bed. It is believed to be fast acting and to have the potential to reduce common sleep problems, such as, difficulty getting to sleep, waking continually throughout the night, and waking early in the morning. If you choose to participate in this study, you may be asked to take one SX3752 tablet each night for one week, or you may be assigned to a no-treatment control group, in which case you would not actually receive any SX3752 tablets.

This study is being conducted by:
Prof. Phyllis Butow (Chief Investigator),
University of Sydney
Prof. Robert Boakes,
University of Sydney
Mr. Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to:

Initial Session (Approximately 30min)
1) Provide some personal details, e.g. age, gender.
2) Complete a questionnaire about your sleeping habits over the past month.
3) Complete a questionnaire about some of your views on medical treatment.

Over the next week
1) Take one SX3752 tablet each night before you go to bed for the next week. You take these tablets by mouth with a glass of water.

At the end of the week (Take-home Questionnaires, Approximately 30min)
1) Complete a questionnaire about the quality of your sleep over the week.
2) Complete a questionnaire about how effective you think the SX3752 tablets are.
3) Complete a questionnaire about your general health symptoms over the week.

Return these questionnaires within one week of the final day of your treatment.
Introduction
You are invited to take part in a research study investigating the effectiveness of a new herbal medication, SX3752, designed to reduce difficulty sleeping. Difficulty sleeping is one of the most common general health problems reported today. It includes taking a long time to get to sleep, waking up frequently, and waking up earlier than expected and not being able to return to sleep. The main purpose of this study is to assess the effectiveness of SX3752 for reducing difficulty sleeping over a one week period. SX3752 is a pill that you take by mouth with a glass of water each night before you go to bed. It is believed to be fast acting and to have the potential to reduce common sleep problems, such as, difficulty getting to sleep, waking continually throughout the night, and waking early in the morning. If you choose to participate in this study, you may be asked to take one SX3752 tablet each night for one week, or you may be assigned to a no-treatment control group, in which case you would not actually receive any SX3752 tablets.

This study is being conducted by:
Prof. Phyllis Butow (Chief Investigator),
University of Sydney
Prof. Robert Boakes,
University of Sydney
Mr. Ben Colagiuri,
University of Sydney

Study Procedures
If you agree to participate in this study, you will be asked to sign the Participant Consent Form. You will then be asked to:

Initial Session (Approximately 30min)
1) Provide some personal details, e.g. age, gender.
2) Complete a questionnaire about your sleeping habits over the past month.
3) Complete a questionnaire about some of your views on medical treatment.

Over the next week
1) Take one SX3752 tablet each night before you go to bed for the next week. You take these tablets by mouth with a glass of water.

At the end of the week (Take-home Questionnaires, Approximately 30min)
1) Complete a questionnaire about the quality of your sleep over the week.
2) Complete a questionnaire about how effective you think the SX3752 tablets are.
3) Complete a questionnaire about your general health symptoms over the week.
4) Return these questionnaires within one week of the final day of your treatment.
**Risks**
Some people who take SX3752 report experiencing a decreased appetite in response to the medication. If you do experience a decreased appetite while taking SX3752 this will go away completely when you stop taking the tablets.

**Benefits**
While we intend that this research study furthers medical knowledge and may help people experiencing difficulty sleeping, it may not be of direct benefit to you.

**Costs**
Participation in this study will not cost you anything, nor will you be paid.

**Voluntary Participation**
Participation in this study is entirely voluntary. You do not have to take part in it. If you do take part, you can withdraw at any time without having to give a reason. Whatever your decision, please be assured that it will not have any repercussions.

**Confidentiality**
All the information collected from you for the study will be treated confidentially, and only the researchers named above will have access to it. The study results may be presented at a conference or in a scientific publication, but individual participants will not be identifiable in such a presentation.

**Further Information**
When you have read this information, Ben Colagiuri will discuss it with you further and answer any questions you may have. If you would like to know more at any stage, please feel free to contact him on 0402 236 400, or by email at benc@psych.usyd.edu.au.

Any person with concerns or complaints about the conduct of a research study can contact the Senior Ethics Officer, Ethics Administration, University of Sydney, on (02) 9351 4811 (Telephone); (02) 9351 6706 (Facsimile) or gbrody@mail.usyd.edu.au (Email).
Appendix 4j – Posttreatment Questionnaire for Study 3, Experiment 3

Post Treatment Questionnaire

Instructions: Please answer the following questions about your treatment with SX3752 over the last week as honestly as possible.

1. How effective were the SX3752 tablets for reducing your sleep difficulty?
   NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  VERY EFFECTIVE

2. At the beginning of the study, how effective did you expect the SX3752 tablets would be for reducing your sleep difficulty?
   NOT AT ALL  0  1  2  3  4  5  6  7  8  9  10  VERY EFFECTIVE

3. Did you take one SX3752 tablet every night?  YES  NO
   If no, how many nights did you miss taking the SX3752 tablet________
   What led to you not taking the SX3752 tablets?
   ____________________________________________________________
   ____________________________________________________________

4. Did you experience any side effects while you were taking the SX3752 tablets?  YES  NO
   If yes, what were these side effects and how severe were they?
   Side effect 1 (please specify) _______________________________
   VERY MILD  0  1  2  3  4  5  6  7  8  9  10  VERY SEVERE
   Side effect 2 (please specify) _______________________________
   VERY MILD  0  1  2  3  4  5  6  7  8  9  10  VERY SEVERE
   Side effect 3 (please specify) _______________________________
   VERY MILD  0  1  2  3  4  5  6  7  8  9  10  VERY SEVERE
Appendix 4k – Adverse Symptom Checklist for Study 3, Experiment 3

General Health Questionnaire

Instructions:
Please answer the following questions about your general health over the last week as honestly as possible.

1. During the last week, have you felt more tired than usual when you wake up? YES  NO
   If yes, on how many days did this happen _________
   On average, how much more sleepy did you feel than usual?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 VERY MUCH MORE

2. During the last week, have you felt more tired than usual in the evenings? YES  NO
   If yes, on how many days did this happen _________
   On average, how much more tired did you feel than usual?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 VERY MUCH MORE

3. During the last week, have you experienced any headaches? YES  NO
   If yes, on how many days did you experience headaches _________
   On average, how severe would you rate these headaches?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

4. During the last week, have you experienced any dizziness? YES  NO
   If yes, on how many days did you feel dizzy _________
   On average, how severe would you rate these dizziness?
   NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

5. During the last week, have you experienced any nausea? YES  NO
6. During the last week, have you experienced a dry mouth? YES NO

If yes, on how many days did you have a dry mouth________

On average, how severe would you rate this dry mouth?
NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

7. During the last week, have you noticed any changes in your mood? YES NO

If yes, please describe these changes in your mood:

_____________________________________________________

_____________________________________________________

_____________________________________________________

On average, how severe would you rate this change in mood?

NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE

8. During the last week, have you noticed any changes in your appetite? YES NO

If yes, please describe these changes in your appetite:

_____________________________________________________

_____________________________________________________

_____________________________________________________

On average, how severe would you rate this change in appetite?

NOT AT ALL 0 1 2 3 4 5 6 7 8 9 10 EXTREMELY SEVERE
9. During the last week, have you noticed any changes in your energy levels?  YES  NO

If yes, please describe these changes to your energy levels:

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

On average, how severe would you rate this change in energy level?

NOT AT ALL  0     1     2     3     4     5     6     7     8     9     10     EXTREMELY SEVERE

10. Do you have any other comments about your general health over the past week?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

END OF QUESTIONNAIRE