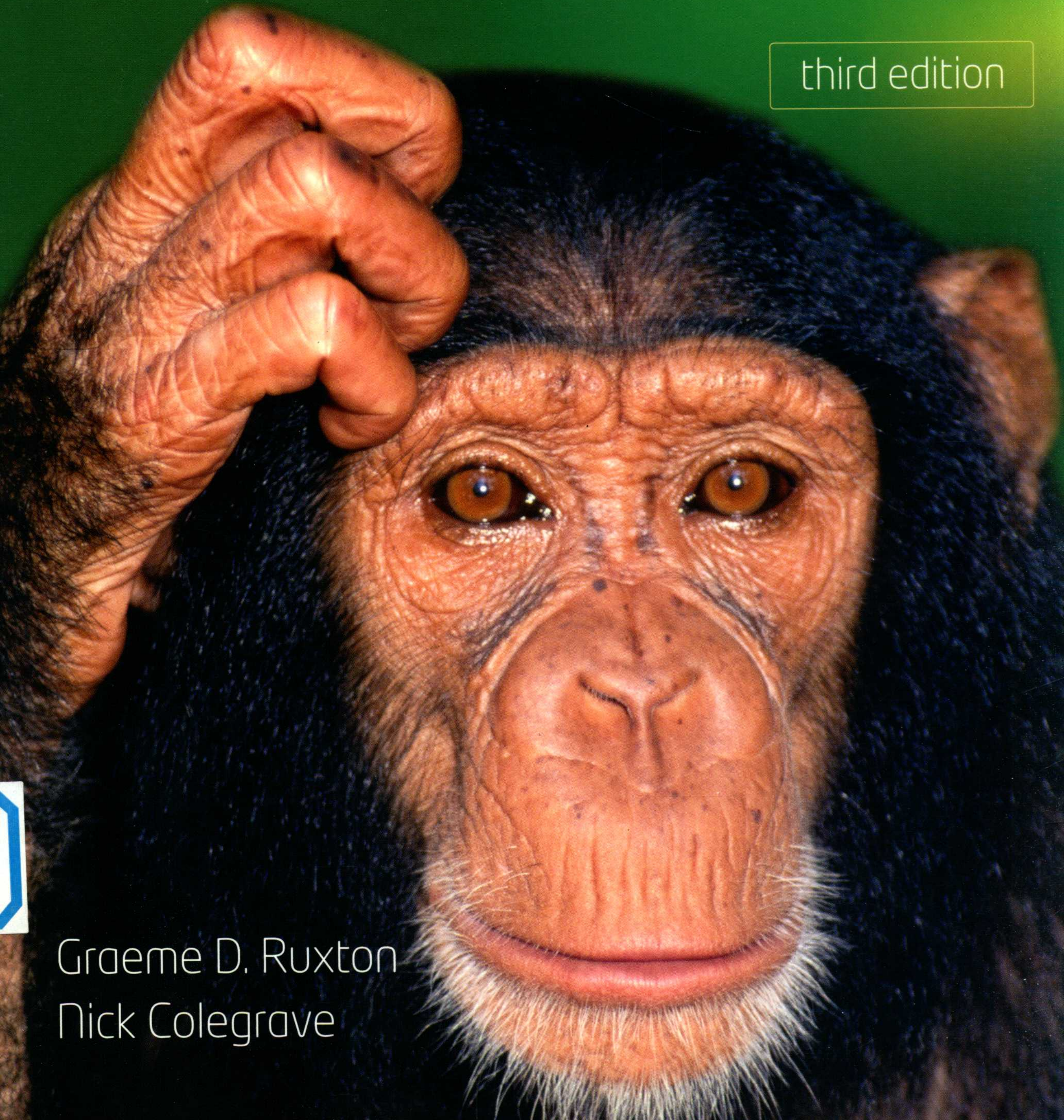
OXFORD

experimental design for the life sciences

third edition



Graeme D. Ruxton
Nick Colegrave

Experimental design for the life sciences

THIRD EDITION

Graeme D. Ruxton

University of Glasgow

Nick Colegrave

University of Edinburgh

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan South Korea Poland Portugal
Singapore Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States
by Oxford University Press Inc., New York

First edition 2003
Second edition 2006

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by MPS Limited, a Macmillan Company
Printed in Great Britain on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire

ISBN 978-0-19-956912-0

1 3 5 7 9 10 8 6 4 2

Experimental design for the life sciences

To Hazel and Becky

Preface

How to read this book

This book is an introduction to experimental design. We mean it to be a good starter if you have never thought about experimental design before, and a good tune-up if you feel the need to take design more seriously. It does not come close to being the last word on experimental design. We cover few areas of design exhaustively, and some areas not at all. We use the Bibliography to recommend some good books that would facilitate a deeper consideration of design issues. That said, it is also important to realise that the basic ideas of design covered in this book are enough to carry out a wide range of scientific investigations. Many scientists forge a very successful career using experiments that never take them outside the confines of the material covered in this book. This book will also help you tackle more advanced texts, but if you absorb all that we discuss here, then you may find that you know all the experimental design that you feel you need to know.

This book is about how to design experiments so as to collect good quality data. Of course, those data will almost certainly need statistical analysis in order to answer the research questions that you are interested in. In order to keep the size of this book manageable, we do not enter into the details of statistical analysis. Fortunately, there are a huge number of books on statistical analysis available: we even like some of them! We recommend some of our favourites in the Bibliography. We also provide some pointers in the text to the types of statistical tests that different designs are likely to lead to.

We often illustrate our points through use of examples. Indeed, in some cases we found that the only way to discuss some issues was through examples. In other cases, our point can be stated in general terms. Here we still use examples to amplify and illustrate. Although we think that you should aim to read the book right through like a novel at least once, we have tried to organize the text to make dipping into the book easy too.

Our first job will be to remove any doubt from your mind that experimental design is important, and so we tackle this in Chapter 1. Chapter 2 discusses how good designs flow naturally from clearly stated scientific questions. Almost all experiments involve studying a sample and extrapolating conclusions about that sample more widely; how to select a good sample is the key theme of Chapter 3. The nitty-gritty of some simple designs forms the basis of Chapter 4. There is no point in formally designing an elegant experiment if you then make a poor job of actually collecting data from your sample individuals, so Chapter 5 is about some tips on taking effective measurements. Chapter 6 is a compendium of slightly more specialized points, which do not fit naturally into the other chapters but at least some of which ought be useful to your branch of the life sciences.

This is *not* a maths book. Nowhere in this book will you find the equation for the normal distribution, or any other equation for that matter. Experimental design is not a subsection of maths, you don't need to be a mathematician to understand simple but effective designs, and you certainly don't need to know maths to understand this book.

Good experimental design is vital to good science. It is generally nothing like as difficult as some would have you believe: you can go a long way with just the few simple guidelines covered in this book. Perhaps most amazingly of all, it is possible to derive enjoyment from thinking about experimental design: why else would we have wanted to write this book!

On the second edition

The coverage has not changed dramatically from the first edition, although all sections have been rewritten—sometimes extensively so—for greater clarity. We have also attended to filling in a few gaps, most notably in substantially increasing our consideration of the special challenges associated with using human subjects. However, the big change in this edition is the way in which the material is presented. As well as extending the use of some of the features of the previous edition, we also include a number of new features that we hope will substantially increase the ease of use of this book. We introduce these in the following pages.

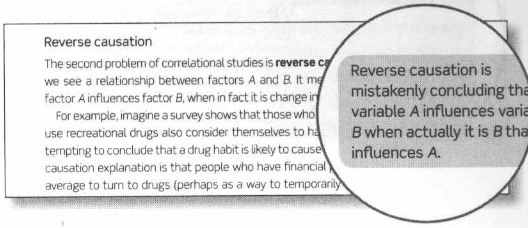
On the third edition

We have very much retained the format of the second section and cover much of the same material. However, we have again taken advantage of feedback to rewrite sections for improved clarity. This should be especially apparent in Chapter 3, where we have extensively revised and expanded our coverage of the key subjects of between-individual variation, replication, and statistical power, using an entirely new example for illustration throughout the chapter. We have also tried to rectify some omissions from the previous editions. In particular, we have always discussed in Chapter 2 how to go from a biological question to hypotheses, and then the design of an experiment. However, in previous editions, we had ducked the issue of where questions come from in the first place. Since this is an issue that early-career scientists in particular have challenged us to address, we now offer some general guidance on this issue. Another notable inclusion in this edition is greater consideration of human biomedical studies. In particular, we have added sections on how correlative studies can be particularly important in studying human epidemiology and on the specialist terminology used in the biomedical literature to describe experimental designs. We have also increased our discussion of the advantages of Latin square designs and increased our coverage of how to make measurements (for example, introducing the concepts of sensitivity and specificity of a test). Finally, we have increased our use of figures throughout, and have greatly enhanced the captions of each figure to allow figures to be understood without consultation of the surrounding text.

Learning features

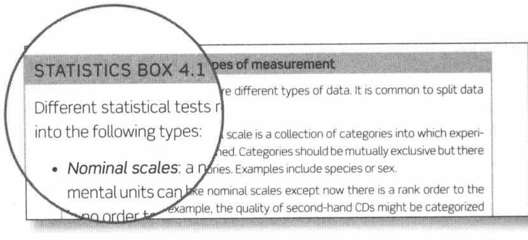
Key definitions

There is a lot of jargon associated with experimental design and statistical analysis. We have not tried to avoid this. Indeed, we have deliberately tried to introduce you to as much of the jargon as we can. By being exposed to this jargon, we hope that it will become second nature to you to use—and understand—it. This should also make reading more advanced texts and the primary literature less daunting. However, to make negotiating the minefield of jargon more straightforward, we have increased the number of definitions of key terms provided and increased the level of detail in all definitions. Each key word or phrase is emboldened at the point where we first use it, and is given a clear definition in a box nearby.



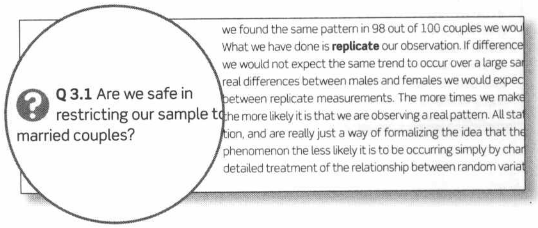
Statistics boxes

To emphasize the important link between good experimental design and statistics we now include a number of statistical boxes. These boxes should help you to see how thinking about design helps you think about statistics and vice versa. We have added the boxes at points where we think that keeping the statistics in mind is particularly useful and include pointers in the main text to each box.



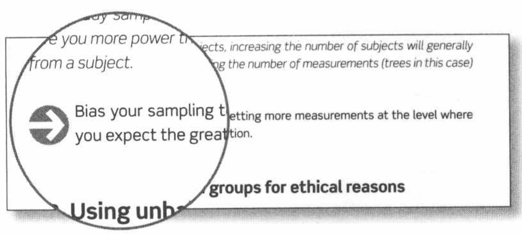
Self-test questions

The more you think about experimental design, the easier designing robust experiments becomes. To get you thinking while you are reading this book, we now include a number of self-test questions in every chapter. Often there will not be a clear right or wrong answer to a question, but suggested answers to all questions can be found at the back of the book.

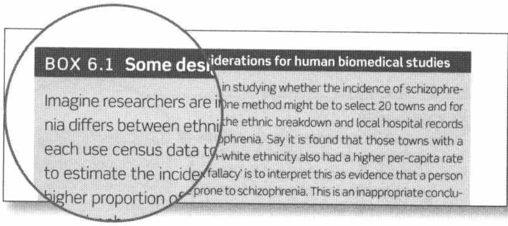


Take-home messages

To help to consolidate your thinking, we end most sections with a take-home message. By including these throughout the text we hope to give you an immediate opportunity to evaluate your understanding of a section before moving on.



Boxes



We have aimed to produce a book that can be read from cover to cover, and so have tried to keep unnecessary details out of the main text. However, in areas where we feel that more details or more examples will lead to a fuller understanding of a concept we have included supplementary boxes.

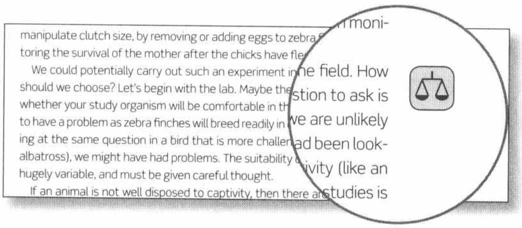
Chapter outlines

Every chapter begins with an outline of the main points covered by that chapter. These outlines should help you to prepare for what is ahead. The outlines will also provide an easy way for you to dip in and out of the book during subsequent readings.

Chapter summaries

Every chapter finishes with a summary of the most important points covered in that chapter. By reading through this list you should be able to reassure yourself that you have got everything out of the chapter that you can, or go back and read again sections that are not yet clear.

Ethical issues



We cannot overemphasize the importance of ethical issues when designing studies involving living organisms (including humans) or material derived from living organisms. In this new edition we highlight sections of particular relevance to ethics by placing a symbol in the margin at the appropriate point.

Flow chart

The exact process of designing an experiment will vary considerably between studies. Nevertheless, there are key stages in the design process that will apply to most, if not all, studies. The flow chart at the end of this book is intended to summarize and guide you through the main stages of designing an experiment. We have indicated at each point in the chart the sections of the book that are most relevant.

Acknowledgements from the first edition

Graeme has been moaning at the representatives of scientific publishers for years that there was not a book like this available. However, it was Sarah Shannon of Blackwell's that suggested that he stop moaning and write one himself. It was also Sarah who suggested that Nick would be an ideal co-author. We may have ended up at another publisher, but we are both very grateful to Sarah for her support and advice at the inception of this book.

At OUP, Jonathan Crowe has been everything we could have hoped for in an editor. He commented on two drafts, and found just the right combination of criticism and enthusiasm to spur us on. He also found excellent reviewers (Roland Hall, Keith McGuiness, and two others), who, in their own diverse ways, have also really helped to improve the finished work. Also at OUP, John Grandidge and Emily Cooke were models of friendly efficiency. Many students from the 2001-2002 BSc Zoology and Aquatic Bioscience courses at Glasgow, and the Zoo4 QZ course at Edinburgh provided useful feedback. In particular, Sandie King provided particularly full and useful comments. Fiona McPhie read the whole thing and gave honest feedback on some bad jokes that we've thankfully removed.

The text was carefully proof-read by Hazel Ruxton and the late David Ruxton, who managed to remove embarrassing numbers of errors.

Many people have influenced our thoughts on experimental design over the years, and as a result some of the examples will have been borrowed in whole or part. While it is impossible to trace and thank the originators of all these ideas there are some people we should certainly mention. Graeme learnt a huge amount about designing experiments from Will Cresswell and Rik Smith. Kate Lessells and Andrew Read have both, over the years, forced Nick to think far harder about experiments than he ever really intended; they will both find large elements of themselves throughout the book. Anyone else that recognizes, and wishes to claim as their own, examples contained in this book is welcome to contact us. If the publishers ever let us write a second edition, we will make sure you get the proper credit.

The figures were drawn by the imaginative and efficient Liz Denton. It's been an absolute pleasure working with her. Stuart Humphries kindly provided the drawing of a bird for Figure 2.1. Thanks also to Ruedi Nager, Stuart Humphries, and James Brown of the Scottish Agricultural Science Agency for cover illustrations.

Lastly we'd like to thank each other. Co-writing a book is much easier than working alone. We both found that the other person can really pick you up whenever inspiration disappears and writers' block is looming. But best of all, no matter what mistakes or shortcomings you find in this book, we'll both be able to blame it on the other guy!

Acknowledgements from the third edition

Thanks again to Liz Denton for the figures. At OUP, Dewi Jackson has helped to guide us through the process of putting the new edition together. Thanks to all the students and colleagues who have offered feedback on previous editions. Thanks to Hazel and Becky for still being with us through all the editions. Thanks also to three new additions since we first wrote the book—Katherine, Amelie, and Isla—for enriching our lives beyond measure, and providing happy distraction from these revisions.

Contents

Why you should care about design

1

1.1 Why experiments need to be designed

When life sciences students see the phrase 'experimental design', it can either ease them gently into a deep sleep or cause them to run away screaming. For many, experimental design conjures up unhappy memories of mathematics or statistics lessons, and is generally thought of as something difficult that should be left to statisticians. Wrong on both counts! Designing simple but good experiments doesn't require difficult maths. Instead, experimental design is more about common sense, biological insight, and careful planning. Having said that, it does require a certain type of common sense, and there are some basic rules. In this book, we hope to steer you towards thinking more effectively about designing experiments.

Designing effective experiments needs thinking about biology more than it does mathematical calculations.

So why are many life scientists so averse to thinking about design? Part of the reason is probably that it is easy to think that time spent designing experiments would be better spent actually doing experiments. After all, the argument goes, we are biologists so let's concentrate on the biology and leave the statisticians to worry about the design and analysis. This attitude has given rise to a number of myths that you can hear from the greenest student or the dustiest professor.

Myth 1: ***It does not matter how you collect your data, there will always be a statistical 'fix' that will allow you to analyse it.***

It would be wonderful if this was true, but it is not. There are a large number of statistical tests out there, and this can lead to the false impression that there must be one for every situation. However, all statistical tests make assumptions about your data that must be met before the test can be meaningfully applied. Some of these assumptions are very specific to the particular test. If you cannot meet these, there may be a substitute test that assumes different characteristics of your data. But you may find that this alternative test only allows you to use your data to answer a different scientific question