

**Critical Thinking**  
Understanding and Evaluating Dental Research  
Third Edition





# Critical Thinking

## Understanding and Evaluating Dental Research

Third Edition



**Donald Maxwell Brunette, PhD**

Professor Emeritus  
Department of Oral Biological and Medical Sciences  
Faculty of Dentistry  
University of British Columbia  
Vancouver, Canada

With contributions by

**Ben Balevi, DDS, MSc**  
Adjunct Professor  
Faculty of Medicine  
University of British Columbia  
  
Private Practice  
Vancouver, Canada

**Helen L. Brown, MLIS, MAS, MA**  
Reference Librarian  
University of British Columbia Library  
Vancouver, Canada



QUINTESSENCE PUBLISHING

Berlin, Barcelona, Chicago, Istanbul, London, Mexico City,  
Milan, Moscow, Paris, Prague, São Paulo, Seoul, Tokyo, Warsaw



## Library of Congress Cataloging-in-Publication Data

Names: Brunette, Donald Maxwell, author.

Title: Critical thinking : understanding and evaluating dental research / Donald Maxwell Brunette.

Description: Third edition. | Batavia, IL : Quintessence Publishing Co, Inc, [2019] | Includes bibliographical references and index.

Identifiers: LCCN 2019017538 (print) | LCCN 2019018451 (ebook) | ISBN 9780867158014 (ebook) | ISBN 9780867158007 (softcover)

Subjects: | MESH: Dental Research

Classification: LCC RK80 (ebook) | LCC RK80 (print) | NLM WU 20.5 | DDC 617.6/027--dc23

LC record available at <https://lcn.loc.gov/2019017538>



© 2020 Quintessence Publishing Co, Inc  
Quintessence Publishing Co Inc  
411 N Raddant Road  
Batavia, IL 60510  
[www.quintpub.com](http://www.quintpub.com)

5 4 3 2 1

All rights reserved. This book or any part thereof may not be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, or otherwise, without prior written permission of the publisher.

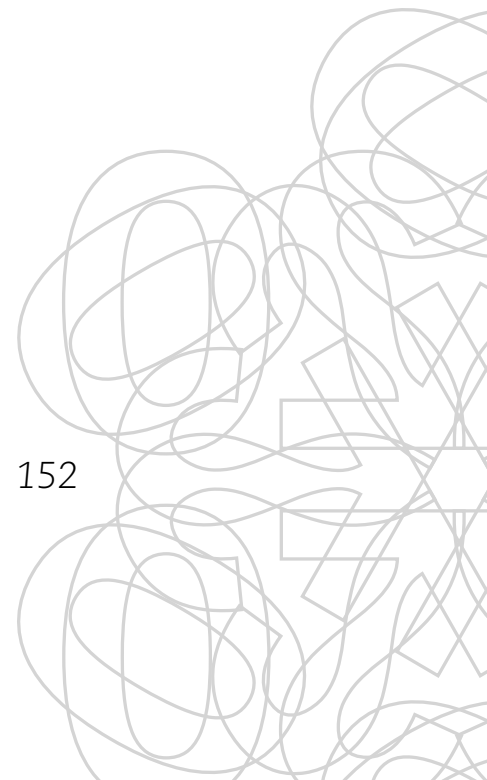
Editor: Zachary Kocanda  
Design: Sue Zubek  
Production: Sarah Minor

Printed in Korea

# Table of contents



	<i>Dedication</i>	vii
	<i>Preface</i>	viii
1	Reasons for Studying Critical Thinking	1
2	Scientific Method and the Behavior of Scientists	13
3	The Components of a Scientific Paper	32
4	Rhetoric	40
5	Searching the Dental Literature with Helen L. Brown	55
6	Logic: The Basics	74
7	Introduction to Abductive and Inductive Logic: Analogy, Models, and Authority	84
8	Causation	97
9	Quacks, Cranks, and Abuses of Logic	110
10	Elements of Probability and Statistics, Part 1: Discrete Variables	128
11	Elements of Probability and Statistics, Part 2: Continuous Variables	145
12	The Importance of Measurement in Dentistry	152





13	Errors of Measurement	163
14	Presentation of Results	173
15	Diagnostic Tools and Testing with Ben Balevi	197
16	Research Strategies and Threats to Validity	232
17	Observation	245
18	Correlation and Association	257
19	Experimentation	274
20	Experiment Design	288
21	Statistics As an Inductive Argument and Other Statistical Concepts	301
22	Judgment	315
23	Introduction to Clinical Decision Making with Ben Balevi	333
24	Exercises in Critical Thinking	356
	<i>Appendices</i>	375
	<i>Index</i>	397

# Dedication



To the late Doug Waterfield, with whom I collaborated extensively once it became clear through the in vivo studies of Babak Chehroudi that macrophages played an important role in tissue response to the topography of implants. Doug was a tireless advocate for students in the Faculty of Dentistry and spared no effort to help them succeed.

# Preface



The need for a third edition of *Critical Thinking* was driven by developments in dental and biomedical research conditions and practice, as well as reader comments. The third edition emphasizes the practical application of critical thinking in the production and evaluation of scientific publications. A number of issues are now more prominent.

Thanks to the Internet, patients of today have much more information (both accurate and inaccurate) than the patients of yesteryear. The consequence to dentists is that they must explain and be prepared to justify their clinical decisions with evidence. This requires that dentists be capable of unearthing relevant information and evaluating alternatives by accepted criteria so that patients can be confident in their dentist's recommendations. In line with the virtual revolution entailed in dealing with the ever-expanding scientific literature, the chapter on searching the dental literature now emphasizes scholarly publishing issues at all stages of research in addition to covering more traditional topics such as types of resources and searching techniques. The preparation of systematic reviews is discussed in greater detail, as are the strengths and weaknesses of various services such as Web of Science and Google Scholar as well as approaches to identifying relevant evidence in grey literature and its use in reviews. The author of the revised chapter is Helen Brown, who has taken over from Kathy Hornby, for whose contribution I continue to be grateful. Helen has master's degrees in English literature and library science, along with frontline experience resolving student problems in literature searches. She has made significant changes in the emphasis of this chapter. For example, the increasing trend toward open access publications and open science initiatives is addressed, providing important information for both prospective authors and readers with limited access to expensive subscription-based resources. The strengths and weaknesses of the h-index, now widely used

by researchers to demonstrate the impact of their work, is also now considered, and there is updated information on impact factors of dental journals and other bibliographic tools.

A second new contributor is Ben Balevi, who takes over the chapter on diagnostic tests and measures from my former student Carol Oakley, whose contributions continue to influence the current chapter. Ben has a master's degree in evidence-based dentistry and has been recognized for his contributions in this area by the American Dental Association.

There have been considerable changes in other aspects of the book as well. A new chapter discusses the function of each section of the scientific paper so that researchers can check that their papers are fulfilling the expectations of readers and referees.

The coverage on logic has expanded to include more material on informal logic, for example, the use of Walton's critical questions for evaluating causation. The informal logic used in the refereeing process in which the burden of proof shifts from authors to reviewers and back has been added to complement the argumentation maps found in the chapter on judgment that features tactical details such as rebuttals, caveats, and supporting evidence. The intent is to give investigators a practical approach to anticipating and dealing with criticisms. The chapter on judgment now includes materials of both historical (Darwin's decision to marry and Benjamin Franklin's moral algebra) and novel decision-making heuristics, such as fast frugal heuristics.

Clinical decision making has been given its own chapter and expanded to include detailed, worked-out calculations and clinical scenarios. In addition to decision tree analysis, two elements of central importance to patients are presented with worked-out calculations: cost-effectiveness analysis and willingness to pay analysis.

The chapter on diagnostic tools and testing has been expanded to include new worked-out calculations and examples, as well as illustrations



to clarify this topic, which many find difficult and confusing.

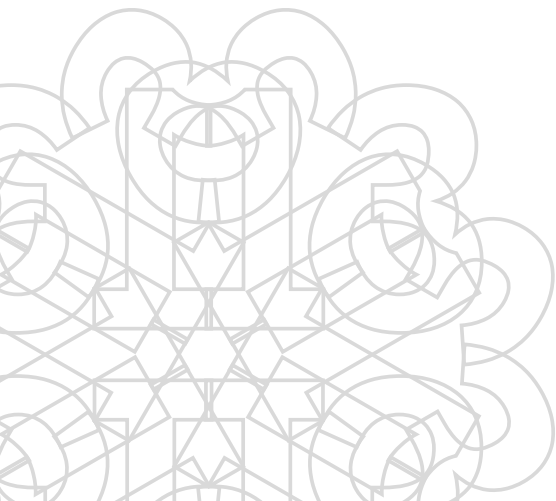
The chapter on experimentation now covers randomized field trials, a method employed in causally dense environments such as business problems, often combined with trial-and-error approaches. The technique is evolving but has been found useful by organizations such as Google that run thousands of randomized field trials.

The statistics chapter has expanded coverage of Bayesian statistics, as this approach is expected to become more widely used. The chapter on presenting results contains new illustrations and more extensive discussion of the integrity-based principles for displaying information emphasized by Tufte. The section on the behavior of scientists has been expanded with discussion on the reproducibility crisis now shaking some scientists' faith in the reliability of reported results, even including those published in high-impact journals. All chapters have also been improved by the addition of updates, revisions, or examples.

I acknowledge the support of my colleagues who have never hesitated to provide advice in their respective domains of expertise. I also thank my former students, friends, and collaborators, whose insights

inform numerous sections. Regular University of British Columbia (UBC) contributors in this way include Babak Chehroudi, Tim Gould, Hugh Kim, Chris Overall, Mandana Nematollahi, Eli Konorti, and Nick Jaeger (Electrical and Computer Engineering). Colleagues outside UBC, Christopher McCulloch (University of Toronto), Ken Yaegaki (Nippon Dental University), and Doug Hamilton (Western University), have also provided advice and encouragement. I also thank my collaborators at the Swiss Institute of Technology (Zurich), Professors Marcus Textor and Nicolas Spencer, and my former post-doc Dr Marco Wieland, for continuing my education in surface science. Of course, I would be an ingrate if I failed to acknowledge the continuing support of my wife Liz, sons Max and Regan (and their partners Heather and Malizza), who provided various examples used in the book including the exploits of my problem-solving grandson Calixte and developing gymnast granddaughter Genèveve.

I also thank Quintessence for their ongoing support through three editions of this work. This edition has been facilitated through the efforts of Bill Hartman and Bryn Grisham. Zach Kocanda has been exemplary in his attention to detail, often under trying circumstances, in editing the manuscript.





# 1

## Reasons for Studying Critical Thinking

### Critical Thinking

*Critical thinking* has been defined many ways, from the simple—“Critical thinking is deciding rationally what to or what not to believe”<sup>2</sup>—to the more detailed “Critical thinking is concerned with reason, intellectual honesty, and open-mindedness, as opposed to emotionalism, intellectual laziness, and closed-mindedness”<sup>3</sup>—to the nearly comprehensive:

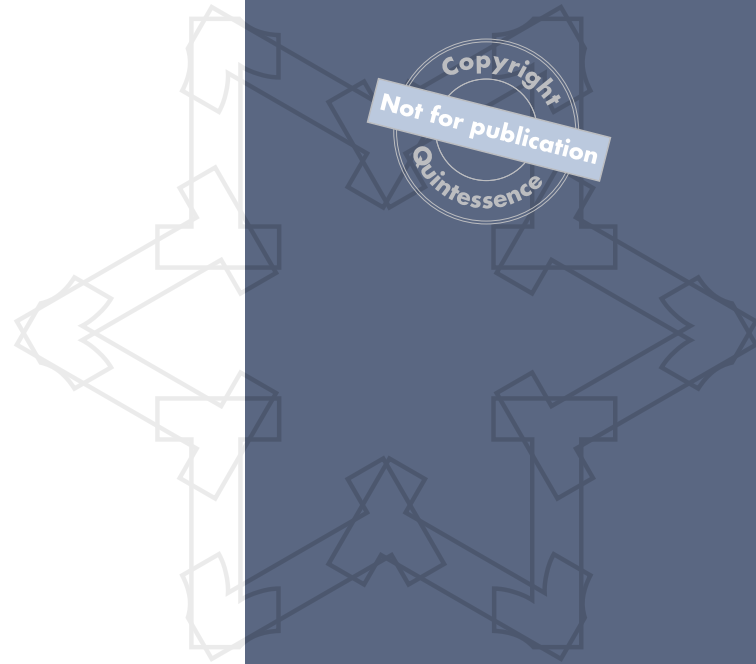
Critical thinking involves following evidence where it leads; considering all possibilities; relying on reason rather than emotion; being precise; considering a variety of possible viewpoints and explanations; weighing the effects of motives and biases; being concerned more with finding the truth than with being right; not rejecting unpopular views out of hand; being aware of one’s own prejudices and biases; and not allowing them to sway one’s judgment.<sup>3</sup>

Self-described practitioners of critical thinking range from doctrinaire postmodernists who view the logic of science with its “grand narratives” as inherently subordinating<sup>4</sup> to market-driven dentists contemplating the purchase of a digital impression scanner. In this book, critical thinking, and in particular the evaluation of scientific information, is conceived as “organized common sense” following Bronowski’s view of science in general.<sup>5</sup> Of course, common sense can be quite uncommon. A secondary use of the term *critical thinking* implies that common sense involves a set of unexamined and erroneous assumptions. For example, prior to Galileo, everyone “knew” that heavy objects fell faster than lighter ones. Critical thinking as organized common sense takes the systematic approach of examining assumptions. The professional use of critical thinking is particularly complex for dental professionals because they live in two different worlds. On the one hand, they are health professionals treating patients who suffer from oral diseases.



**It has happened more than once that I found it necessary to say of one or another eminent colleague, ‘He is a very busy man and half of what he publishes is true but I don’t know which half.’”**

ERWIN CHARGAFF<sup>1</sup>





## REASONS FOR STUDYING CRITICAL THINKING

On the other hand, dentists typically also inhabit the business world, where decisions may be based on the principle of maximizing income from their investment. Dental practice is based only very loosely on responding to disease<sup>6</sup>; less than one-third of patient visits result in identifying a need for restorative care.<sup>7</sup> Twenty percent of work is elective, such as most of orthodontics, tooth whitening, and veneers, and typically that work comprises the most lucrative aspects of practice. Thus, the information that must be evaluated in performing these disparate roles covers the spectrum from advertisements to financial reports to systematic meta-analysis of health research.

Dentists are health professionals, people with specialized training in the delivery of scientifically sound health services. The undergraduate dental curriculum is designed to give dental students the basic knowledge to practice dentistry scientifically, at least to the extent allowed by the current state of knowledge. But if any guarantee can be made to dental students, it is that dentistry will change, because the knowledge base of biomedical and biomaterial sciences grows continually. Most dentists today have had to learn techniques and principles that were not yet known when they were in dental school. In the future, as the pace of technologic innovation continues to increase and the pattern of dental diseases shifts, the need to keep up-to-date will be even more pressing. Means of staying current include interacting with colleagues, reading the dental literature, and attending continuing education courses—activities that require dentists to evaluate information. Yet, there is abundant historical evidence that dentists have not properly evaluated information. Perhaps the best documented example in dentistry of a widely accepted yet erroneous hypothesis is the focal infection theory. Proposed in 1904 and accepted by some clinicians until the Second World War, this untested theory resulted in the extraction of millions of sound teeth.<sup>8</sup> But errors are not restricted to the past; controversial topics exist in dentistry today because new products or techniques are continually introduced and their usefulness debated. Ideally, dentists should become sophisticated consumers of research who can distinguish between good and bad research and know when to suspend judgment. This goal is different from proposing that dentists become research workers. One objective of this book is to provide the skills enabling a systematic method for the evaluation of scientific papers and presentations.

A marked addition to the challenges of dental practice in recent years is that patients have increased access through the Internet to information as well as

misinformation. Dentists thus are more likely to be questioned by patients on proposed treatments, risks, and options. In responding to such questions, it is clearly advantageous for dentists to be able to present a rational basis for their choices. Chapter 23 covers an evidence-based approach to clinical decision making and appendix 9 provides a template for dental offices to use in documenting their decisions based on recent evidence.

A systematic approach to analyzing scientific papers has to be studied, because this activity requires more rigor than the reasoning used in everyday life. Faced with an overabundance of information and limited time, most of us adopt what is called a *make-sense epistemology*. The truth test of this epistemology or theory of knowledge is whether propositions make superficial sense.<sup>9</sup> This approach minimizes the cognitive load and often works well for day-to-day short-term decision making. In 1949, Zipf of Harvard University published *Human Behaviour and the Principle of Least Effort*, in which he stated:

The Principle of Least Effort means, for example, that in solving his immediate problems he will view these against a background of his probable future problems, *as estimated by himself*. Moreover, he will strive to solve his problems in such a way as to minimize the *total work* that he must expend in solving *both* his immediate problems and his probable future problems.<sup>10</sup>

Zipf used data from diverse sources ranging from word frequencies to sensory sampling to support his thesis. Although the methods and style of psychological research have changed, some more recent discoveries, such as the concept of *cognitive miser* in studies of persuasion,<sup>11</sup> coincide with Zipf's principle. Kahneman in *Thinking, Fast and Slow* has elevated the principle to a law noting that we “conduct our mental lives by the law of least effort.”<sup>12</sup>

In science, the objective is not to make easy short-term decisions but rather to explain the phenomena of the physical world. The goal is accuracy, not necessarily speed, and different, more sophisticated, more rigorous approaches are required. Perkins et al<sup>9</sup> have characterized the ideal skilled reasoner as a critical epistemologist who can challenge and elaborate hypothetical models. Where the makes-sense epistemologist or naive reasoner asks only that a given explanation or model makes intuitive sense, the critical epistemologist moves beyond that stage and asks why a model may be inadequate. That is, when evaluating and explaining,

**Table 1-1** | Level of evidence guideline recommendations of the United States Agency for Healthcare Research and Quality

Level	Type of study	Grade recommendation
1	Supportive evidence from well-conducted RCTs that include 100 patients or more	A
2	Supportive evidence from well-conducted RCTs that include fewer than 100 patients	A
3	Supportive evidence from well-conducted cohort studies	A
4	Supportive evidence from well-conducted case-control studies	B
5	Supportive evidence from poorly controlled or uncontrolled studies	B
6	Conflicting evidence with the weight of evidence supporting the recommendation	B
7	Expert opinion	C

RCTs, randomized controlled trials.

the critical epistemologist asks both *why* and *why not* a postulated model may work. The critical epistemologist arrives at models of reality, using practical tactics and skills and drawing upon a large repertoire of logical and heuristic methods.<sup>9</sup>

Psychologic studies have indicated that everyday cognition comprises two sets of mental processes, System 1 and System 2, which work in concert, but there is some debate whether they operate in a parallel or sequential manner. System 1 operates quickly and effortlessly, whereas System 2 is deliberate and requires attention and effort.<sup>12</sup> System 2 is a rule-based system, and engaging System 2 is the surest route to fallacy-free reasoning.<sup>13</sup> System 2 becomes engaged when it catches an error made by the intuitive System 1. The good news is that extensive work by Nisbett and colleagues (briefly reviewed by Risen and Gilovich<sup>13</sup>) showed that people can be trained to be better reasoners and that people with statistical backgrounds were less likely to commit logical fallacies. Nisbett and colleagues further demonstrated that even very brief training was effective in substantially reducing logical and statistical errors. Thus this book has chapters on logic and statistics.

A second objective of the book is to inculcate the habits of thought of the critical epistemologist in readers concerned with dental science and clinical dentistry.

### The scope of the problem

In brief, the problems facing anyone wishing to keep up with developments in dentistry or other health professions are that (1) there is a huge amount of literature, (2) it is growing fast, (3) much of it is useless in terms

of influencing future research (less than 25% of all papers will be cited 10 times in all eternity,<sup>14</sup> and a large number are never cited at all), and (4) a good deal of the research on a clinical problem may be irrelevant to a particular patient's complaint

The actual rate of growth of the scientific literature has been estimated to be 7% per year of the extant literature, which in 1976 comprised close to 7.5 million items.<sup>11</sup> This rate of growth means that the biomedical literature doubles every 10 years. In dentistry, there are about 500 journals available today.<sup>15</sup> Many dental articles are found in low-impact journals, but, ignoring these, there were still 2,401 articles published in 1980 in the 30 core journals.<sup>16</sup> More recently, it has been estimated that about 43,000 dental-related articles are published per year.

However, the problem is not intractable. Relman,<sup>17</sup> a former editor of the *New England Journal of Medicine*, believes that most of the important business of scientific communication in medicine is conducted in a very small sector of top-quality journals. The average practitioner needs to read only a few well-chosen periodicals.<sup>17</sup> The key to dealing with the problem of the information explosion is in choosing what to read and learning to evaluate the information.

Dentists are exposed to diverse information sources, and the important issues vary depending on the source. For example, a dentist may wish to determine whether potassium nitrate toothpastes reduce dentin hypersensitivity. One approach would be to look up a systematic review on this topic in the Cochrane Library,<sup>18</sup> which many regard as the highest level in the hierarchy of evidence (Table 1-1). The skills required to understand the review would include a basic knowledge of statistics and research design. The same dentist, facing



## REASONS FOR STUDYING CRITICAL THINKING

the competitive pressures of his or her local market, might also want to determine whether a particular laser-bleaching process should be adopted for the practice. In that instance, there might not be a relevant Cochrane review, and there may not even be a relevant paper in a refereed journal to support a decision. Available evidence might consist of advertising brochures and anecdotes of colleagues. The dentist may have to employ a different set of skills, ranging from evaluating the lie factor in graphics (see chapter 14) to disentangling rhetoric from fact. Advertisements and salesmanship are persuasive exercises; the chapter on rhetoric (chapter 4) deals with means of persuasion.

Typically, dentists acquire information on innovative procedures through participation in networks in which their colleagues supply informal data on the effectiveness of the innovations. Nevertheless, dentists cite reading peer-reviewed dental literature and experimental studies as the gold standard for determining the quality of innovations.<sup>19</sup> New technology is often introduced into their practices through trial and error; dentists take the pragmatic approach of directly determining what works in their hands in their practice.<sup>19</sup> Doubtless, some of the personal and financial expenses typical of the trial-and-error approach could be reduced with more effective evaluation of information prior to selecting a material or technique for testing.

This book focuses on evaluating refereed scientific papers, but many of the issues of informational quality and questions that should be asked apply equally to other less formal channels of communication.

### What is a scientific paper?

The Council of Biology Editors defines a *scientific paper* as follows:

An acceptable primary scientific publication must be the first disclosure containing sufficient information to enable peers (1) to assess observations; (2) to repeat experiments; and (3) to evaluate intellectual processes; moreover, it must be sensible to sensory perception, essentially permanent, available to the scientific community without restriction, and available for regular screening by one or more of the major recognized secondary services.<sup>20</sup>

Similar ideas were stated more succinctly by DeBakey,<sup>21</sup> who noted that the contents of an article should be new, true, important, and comprehensible.

A good deal of the literature circulated among dentists does not meet these requirements. But even the throwaway or controlled-circulation magazines are little more than vehicles for advertisements; the amount of information published annually appears formidable.

One approach to dealing with a large number of papers is to disregard original papers and receive information secondhand. Dental and medical journals present reviews of current research in specific clinical or scientific fields; some journals, such as *Dental Clinics of North America* and the *Journal of Evidence Based Dentistry*, are exclusively devoted to this approach. Although this tactic reduces the volume of literature to be covered, it does not solve the problem of evaluating the information contained in the reviews. To perform this task effectively, a researcher must be able to assess the soundness of the reviewer's conclusions. In deciding to accept information secondhand, the researcher is also deciding whether the author of the review is a reliable, objective authority. Thus, the problem of evaluation has been changed, but not eliminated.

This book focuses on the primary literature, where it is hoped that new, true, important, and comprehensible information is published. The systematic review, a relatively new review form, attempts to deal with some of the more glaring problems of traditional reviews and is covered briefly in chapter 5. Although useful for some purposes, the systematic review has its own shortcomings, and the researcher must judge how these affect the conclusions. Journals vary in quality; chapter 5 also discusses bibliometric approaches of ranking journals. In the following section, I present a brief review of how articles get published that may help explain some of this variation.

## The Road to Publication

### The author

The author's goal is to make a significant contribution to the scientific literature: a published paper. To accomplish that goal, the author will have to produce a submission for publication whose contents are new, true, important, and comprehensible. Moreover, the author wants to publish the paper in a journal whose readers will likely find the paper of interest and hopefully be influenced by it. As journals vary in the rigor they demand and the length of papers they accept, the author needs to identify the best journal for his or her purposes.

## Refereed versus nonrefereed journals

The first hurdle faced by an article submitted for publication is an editor's decision on the article's suitability for the journal. Different journals have different audiences, and the editors are the arbiters of topic selection for their journal. Editors can reject papers immediately if they think the material is unsuited to their particular journal.

In some journals, acceptance or rejection hinges solely on the opinion of the editor. However, this method is problematic because informed decisions on some papers can only be made by experts in a particular field. Therefore, as a general rule, the most highly regarded journals ask the opinion of such specialists, called *referees* or *editorial consultants*. Referees attempt to ensure that a submitted paper does not demonstrably deviate from scientific method and the standards of the journal. Whether a journal is refereed can be determined by consulting Ulrich's Periodicals Directory ([ulrichsweb.com](http://ulrichsweb.com)). Editors usually provide referees with an outline of the type of information that they desire from the referee. The criteria for acceptance will necessarily include both objective (eg, obvious errors of fact or logic) and subjective (eg, priority ratings) components. Unfortunately, the task of refereeing is difficult and typically unpaid. Refereeing is often squeezed in among other academic activities, so it should not be surprising that it sometimes is not done well and that referees often disagree.

Studies of the reliability of peer-review ratings are disappointing for readers wanting to keep faith in the peer-review system. Reliability quotients, which can range from 0 (no reliability) to 1 (perfect reliability), for various attributes of papers submitted to a psychology journal<sup>22</sup> follow:

- Probable interest in the problem: 0.07
- Importance of present contribution: 0.28
- Attention to relevant literature: 0.37
- Design and analysis: 0.19
- Style and organization: 0.25
- Succinctness: 0.31
- Recommendation to accept or reject: 0.26

Despite such issues, there is evidence that the review process frequently raises important issues that, when resolved, improve the manuscript substantially.<sup>23</sup>

After consulting with referees, the editor decides whether the paper should be (1) published as is—a comparatively rare event; (2) published after suitable revision; or (3) rejected. Journals reject papers in proportions varying from 0% to greater than 90%.

The literature available to dental health professionals ranges the spectrum of refereed to nonrefereed, from low (or no) rejection rates to high rejection rates. The *Journal of Dental Research*, for example, used to have a 50% rejection rate (Dawes, personal communication, 1990), but that has risen so that 25 years later some 90% of submissions are rejected.<sup>24</sup> Even among high-impact journals, however, there is no guarantee that the referees did a good job. In fact, these considerations only serve to reinforce the view “caveat lector”—let the reader beware.

## Editors and referees

The editor of the journal and the referees are the “gatekeepers” who decide whether a manuscript is accepted. In science the basic rule appears to be something akin to “if it doesn't get published, it doesn't exist.” Thus the rewards in science go to those who publish first, not the first scientist to observe a phenomenon. Obviously, pleasing these gatekeepers is essential to a scientific career.

The editor and the referees are the representatives of the readers of the journal. They protect the readers from wasting their time on obviously erroneous or uninteresting or unsuitable or unoriginal or opaque or trivial submissions. The papers in the journal must be relevant to the readership.

An important part of the editor's job is to protect authors from unjust criticism that can arise from such things as personal animosity between an author and a referee or an attempt by a referee to block publication of a competitor's work. Unfortunately, the scientists who are best able to evaluate a submission may be individuals who can suffer most from its publication, as for example occurs when the referee's own work is “scooped” (ie, published earlier by a competitor).

To justify readers' expenditure of time the paper should address a significant problem or concern and provide a significant amount of information. The length of journal articles varies; some journals publish “letters” rather than full-length papers for interesting but only briefly developed findings. Editors are interested in publishing papers that are likely to be cited in the future or, expressed another way, are building blocks for future research or clinical application. Tacker<sup>25</sup> notes that journals differ in the sophistication of their readership. A general medical journal (eg, *JAMA*) is written at the comprehension level of a third-year medical student, whereas a specialty journal is written for a first- or second-year resident. A scientific



## REASONS FOR STUDYING CRITICAL THINKING

journal should be understandable to third- or fourth-year PhD candidates or above in the general field.

### The editor

The editor decides ultimately whether to accept or reject a submission. As a general rule the editor is an unpaid (or lowly paid) volunteer of distinguished standing in the field covered by the journal. The editor defines the scope of the journal (ie, what subjects are published in it), and if a manuscript falls outside the journal's mandate, it will probably be returned promptly to the author. Similarly, an editor may reject a paper on the grounds that a submission does not advance the field sufficiently or has a low potential for future impact. Such judgments are subjective but nevertheless may need to be made. I call this the "de gustibus" standard after the Latin adage, *De gustibus non est disputandum*: "In matters of taste, there can be no disputes." As the adage indicates, if a decision is made on this basis it will be difficult to persuade the editor to reverse it.

Editors are often responsible for diverse other tasks such as recruiting referees and persuading them to submit their reviews in a timely manner. Some journals have associate editors who oversee submissions in their area of expertise, and the editor must coordinate their activities as well as consult with editorial boards and deal with the various business matters. Despite the importance of their job, editors are not always appreciated by their colleagues, who may resent some decisions. Chernin playfully suggests, "Editors are also the people who separate the wheat from the chaff and frequently publish the chaff."<sup>26</sup>

After the manuscript is accepted by the editor, it may be passed on to a managing editor to take the manuscript through the production and publication process. Day<sup>27</sup> states that editors and managing editors have jobs that are made impossible by the attitudes of authors who submit to their journals. For example, authors might ignore the rules and conventions specified by the journal (eg, the format for the references). Or authors and referees may have irreconcilable views, and the editor may be caught in the middle. Given that the editor's decision could affect the author's career, it is clearly wise not to irritate editors or referees, but rather to make their job in dealing with the submission as easy as possible. That is, authors want the editor to like them, and as has been extensively studied in the psychology literature,<sup>28</sup> liking can be a key factor in persuasion, in this case persuading the editor that the submission should be published.

An indicator of what editors want is provided by the instructions given to referees of journals, often in the form of a checklist or a score sheet that incorporates specific questions for reviews completed online. As an example, I compiled an indicator of some of editors' concerns by simply looking at the instructions sent to me by ten journals. The following characteristics were emphasized:

- 90% (ie, 9/10) concise
- 70% clear
- 70% evaluate by section (eg, introduction, methods)
- 70% adequacy of references
- 60% originality
- 60% adequacy of illustrations
- 50% relationship of conclusions to results

Overall the instructions emphasize economy of expression, ignoring the folk wisdom that "Sermons on brevity and chastity are about equally effective."<sup>26</sup> Nevertheless it is useful for prospective authors to obtain a specific checklist for the journal to which they are submitting so that they can attempt to meet the journal's expectations.

### The referees

The referees are unpaid volunteers; nevertheless they do receive some nonmonetary rewards. They get first access to new information in a field that interests them, and their decisions can influence the direction of that field. On occasion that information may be useful—for example, a submission could contain a reference of which the referee was unaware or a new technique that might be beneficial to the referee's own research, or reading the article might prompt an idea for the referee's future research. Finally, in doing a favor to the editor in refereeing a manuscript, the referee might acquire a store of goodwill that might help when his or her own manuscript is submitted to the journal. (Another well-accepted persuasive factor—reciprocation).<sup>28</sup> Nevertheless, refereeing papers is a low-yield task—the referees' efforts help the editor and those whose papers are published, but the referee typically gets no tangible benefit save the good feeling that comes from doing the right thing. Spending their own time on work for which others will benefit is bound to lead to resentment if those potentially benefitted make the task more difficult than it need be. The applicable golden rule then is to do unto the referees as you would have them do unto you. Make it easy for the referees in the hope they will make it easy for you. In



this spirit then authors should attempt to meet the expectations of referees, in particular not wasting their time. In general, referees expect a scientific writing style characterized by the following qualities:

- **Objectivity.** Data obtained from scientific observation should be capable of being repeated by any competent observer, and the interpretations should be similarly identical among investigators. Expressed another way, in the “Storybook” version of scientific method, there is no room in science for subjective personal data collection and interpretation. Sometimes writers attempt to emphasize their objectivity, and this desire to appear objective can lead to overuse of the passive voice. Of course investigators do have an axe to grind, as they want to be published so that they can reap the rewards of publication—recognition and employment being the chief among these. So a tradition has arisen whereby authors attempt to appear to be objective while being strong advocates for their position. Thus, authors make “Verbal choices . . . that capitalize on a convenient myth . . . reason has subjugated the passions.”<sup>29</sup> In any case, readers have come to expect that scientific writers will present at least a veneer of objectivity (practitioners of qualitative methods might disagree), but readers have other expectations of authors as well.
- **Logic.** Logic not only in organization but in sentence structure and meeting reader expectations.<sup>30</sup>
- **Modesty.** Related to the scientific norm of humility (extravagant claims will attract close and probably critical attention).
- **Clarity.** Scientific writers should follow the common advice to all writers such as avoiding misplaced modifiers, dangling participles, nonparallel constructions, stacked modifiers, etc. (There are numerous books on writing style, such as Strunk and White’s *The Elements of Style*<sup>31</sup> or Zinnser’s *On Writing Well*.<sup>32</sup>)
- **Precision.** Use of precise terminology to avoid confusion and the fallacy of equivocation.
- **Brevity.** To conserve readers’ time.
- **Justified reasoning.** Making the reason for statements clear by referring to data in the paper (eg, “see Figure 1”) or references to the literature.
- **Use of signposts** (eg, “our first objective...,” “our second objective...”), linkage, etc.

Referees typically submit their reports by filling out forms online often accompanied by explanatory remarks in an uploaded text file.

Typically the form starts with what might be called high-level assessments—questions like accept

or reject, priority, overall length of the paper. More detailed points are given in the comments to the authors or the editor or editor’s assistant. The comments to the editor give the referee the possibility to offer frank criticism that might be construed by some authors as being insulting. For example, a referee might comment that the paper is poorly written and needs revision by a native English speaker, and such a comment might be insulting to an author who was in fact a native English speaker.

## Referees versus authors

Typically referees make critical comments on the papers they are reviewing ranging from the easily correctable, such as typographic errors or formatting, through more difficult problems to correct, such as lack of clarity in organization, deficiencies such as inappropriate methodology, or erroneous logic, that lead to unsupported conclusions. Typically the referees will number their comments, and the editor will require the author to address each of them. So in effect the authors and each of the referees enter into a debate presided over by the editor, who might also provide some comments, that might be classed according to the conventions of informal logic or pragmatics as a “persuasion dialogue.”<sup>33</sup> The participants are obligated to give helpful and honest replies to their opponents’ questions. In theory each participant in the dialogue is supposed to use arguments exclusively composed of premises that are commitments of the other participant. But in argumentation, as in life, commitments are notoriously difficult to extract from an opponent, and pretty much the best one can hope for is plausible commitment to an opinion based on reasoned evidence. In conducting the argument the participants are also obligated with a burden of proof, which shifts from one to the other during the dialog. For example, in submitting the paper the author, as proponent, assumes the burden of proof for the conclusions of the paper, and the components of the paper (ie, methods, data, figures, tables, and logic), constitute the means of bearing that burden. Similarly the referee, in making a criticism, assumes the burden of proof of justifying the criticism. This may be done by various means such as citing deficiencies in the evidence in the paper, external scientific evidence (such as previously published papers), or expected standards in the field of study. The editor forwards the referees’ criticisms along with a preliminary decision to the authors who, if they want the submission to proceed to publication, are



## REASONS FOR STUDYING CRITICAL THINKING

expected to bear the burden of proof in responding to the criticisms. This dialogue can be carried over several cycles. Often in my experience, it seems that referees seldom accept or commit to the author's arguments; rather they merely concede by terminating discussion. In science, as in life, it is difficult to say "Sorry, I was wrong." In some instances agreement between the referees and the authors is never achieved, but the issues are clarified to an extent that the editor can make a decision.

The question arises of the logic used by editors in making their decisions. First it should be noted that different types of reasoning employ different standards of proof, and this is not unusual in human affairs. In law for example the standard of proof in criminal cases is "beyond reasonable doubt" whereas civil cases are decided on the "balance of probabilities." Scientific arguments can be complex and may entail various forms of logic ranging from the certainty of deductive logic employed in mathematics to inductive logic, which can deal with calculated probabilities, to informal logic that balances many factors but does not necessarily proceed by strict numeric calculation so that the conclusions are classed qualitatively in terms of their relative plausibility.

Perhaps the reasoning process most employed by editors, who have to make a practical decision, would be the pragmatic model devised by the philosopher Stephen Toulmin<sup>34</sup> (see also chapter 22 for more on argumentation maps), which specifies a system for scientific explanation that includes Claims (such as conclusions in the paper) justified by Evidence and Warrants. A Warrant is the means that connects the Claim to the Evidence; it may be, for example, a scientific principle or a connection established by previous work. An important aspect of Toulmin's approach is that it is field dependent so that appropriate standards are employed for differing types of scientific endeavor. One can see this aspect in action in the scientific literature by observing the content of papers where the rigor of the methods, the quantity of data, or the articulation of the findings differ among different fields of science or among the journals within one field of science. It is the editor who determines the standards of his/her journal, and differences between editors in what they consider important findings or flaws can result in a paper rejected by an editor of one journal being accepted by another one. There are other elements in the Toulmin model, including the Rebuttal arguments that restrict or counter the claim and the Qualifier, which indicates the degree of certainty that the proponent assigns to the Claim (eg unlikely, possibly, highly

probable, or beyond any reasonable doubt) and this feature can hold the key to resolution of the issue. Authors can back off or limit their claims to accommodate for the views of the referees, and the editor can in good conscience publish the article.

### The readers

The end users of the published paper, the *readers*, have been defined as anyone who reads the text "with an intentional search for meaning."<sup>35</sup> Editors and referees are knowledgeable about their fields and like the authors suffer from the problem of familiarity with the assumptions, conventions, and expectations of investigators in their field so that they tend to "fill in" what an author might leave out. General readers, however, differ from the editors and referees in that on average they are less familiar with the research field and may lack information required to understand the submission. Expressed another way, they can't fill in what the author leaves out. As the readers vary widely in their expertise, it falls to the author to determine what they are likely to know (ie, what is common knowledge to everyone in the field), and conversely what the author needs to point out to them. Anything that is novel or unusual needs to be described in detail; for example, investigators may vary from the standard methods in their measurements or calculations, and such changes need to be highlighted and explained.

### Editorial independence

Ideally, the contents of the journal should be independent of economic issues, but this is not necessarily the case. Publication of color illustrations can be prohibitively expensive, and many respected journals are publications of learned societies that operate on lean budgets. *The Journal of Dental Research*, for example, is published and subsidized by the International Association for Dental Research. Such a journal would be expected not to be subject to advertisers' influence. Other journals have a need to generate income, and, in some instances, entire issues appear to be sponsored by a commercial interest. It is not unreasonable to wonder whether the advertiser influenced the editorial content, for "he who pays the piper calls the tune." In recent years, Internet-based journals have arisen that are financed by authors through charges per page. As hard copy of the articles are not produced or distributed, costs are minimal, and the potential for profit is

great. There is thus an incentive for such journals to have a very low (or no) rejection rate, and questionable quality may result.

### Three general questions

A scientific paper is not necessarily an unbiased account of observations; it is more likely an attempt to convince the reader of the truth of a position. As noted by Ziman,<sup>36</sup> it is important to realize that much of the research literature of science is intended, rhetorically, to persuade other scientists of the validity of received opinions. Thus, a reader can expect an author to present his or her data in the most favorable light. Tables, figures, and even calculations may be done so that differences between groups are accentuated and the appearance of error is minimized. A reader's defense as a consumer of this information is an attitude of healthy skepticism. Three general questions a skeptical reader should ask are: Is it new? Is it true? Is it important?<sup>37</sup>

#### Is it new?

A minimum requirement for publication in most instances is that the information is new. However, *new* can be defined in various ways. If a paper using standard histologic techniques reporting the development of teeth in lynx were to be published tomorrow, it might well be new, because, as far as I am aware, the development of lynx teeth has not been described previously. However, it probably would not be new in adding anything to our knowledge of tooth development in general. Such a paper would merely fill in the gaps, however small, in our knowledge. I think that journal editors are fairly lenient in their judgments on what constitutes new information. Kuhn<sup>38</sup> states that one of the reasons why normal puzzle-solving science seems to progress so rapidly is that its practitioners concentrate on problems that do not tax their own lack of ingenuity.

The quality that often distinguishes good scientific papers from the mediocre is originality. Funding agencies are probably better gatekeepers of science in this regard, because an essential criterion for funding is originality. Originality can appear in any component of the research process, including the questions being asked, the methods employed, the research design, or even the interpretation. Because science is a progressive business, approaches that were once original and sufficient can with time become derivative and

deficient. Returning to the example, because scientists have been studying tooth development in lynx for years using standard histologic techniques, there is not much hope that reworking the same approach would provide anything exciting; new methods would be required to bring new insights.

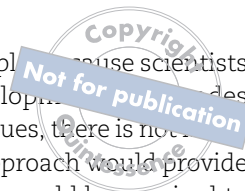
As a consequence of scientific progress, methods become outdated and standards change. Changing standards can be seen in biochemistry by examining the standards for publication of data using polyacrylamide gels. Early publications using the technique showed photographs of gels that did not have good resolution or uniformity and showed poor staining. The photographs of gels were often so uninformative that *Archives of Oral Biology* instructed authors to submit densitometric tracings of the gels. Currently, gel separations are done in two dimensions with superb resolution, and the proteins are stained with much greater sensitivity. A photograph of a gel that would have been acceptable 30 years ago would not be acceptable for publication today. In judging papers, therefore, a key question is whether the techniques and approach are up-to-date as well as whether the question is original.

This principle is so well accepted that investigators sometimes rush to apply new techniques to appear up-to-date. Fisher,<sup>39</sup> the pioneer statistician and author of a classic work on experimental design, warned, "any brilliant achievement . . . may give prestige to the method employed, or to some part of it, even in applications to which it has no special appropriateness."

An exception to the requirement of "newness" for a publication is the need to report confirmations of previous work. One journal for which I refereed placed the category "valuable confirmation of previous work" in third place in their ranking system below exciting original research and interesting new findings, but above categories related to rejection. This type of research is taking on increasing importance in the light of the "reproducibility crisis" to be discussed later.

#### Is it true?

Sound conclusions are the result of reliable observations combined with valid logic. Knowledge of measurement, types of observational errors, experimental design, and controls give some basis to assessments of the reliability of observations. Thus, sections of this book deal with these topics and the logic used to interpret the observations. But the ultimate test of any scientific observation is *reproducibility*; indeed, a practical definition of truth for the purposes of pragmatic working scientists is that a scientific statement is true if





## REASONS FOR STUDYING CRITICAL THINKING

it allows us to make useful, reliable predictions that are confirmed when tested by a competent scientist under the specified conditions. There are theoretical or practical limitations to any approach. Newton's laws of motion would be perfectly valid when applied to billiard balls colliding on a pool table but not useful at very small scales of the world of subatomic particles where quantum physics would be preferred. Note that confirmation does not imply the exact same numeric result, but rather one that is within the specified interval of reported uncertainty.

A clue to the reproducibility of an observation is the consistency of the results within the body of the paper. Another means for evaluating the reliability of observations in a paper is to read what other scientists have written about the work, and citation analysis is an efficient means of uncovering that information. For various reasons, to be discussed later, there is a current "reproducibility crisis" perceived in which confidence in the reproducibility of findings, even those published in high-impact journals, is waning.

A student might wonder whether it is necessary to learn such diverse concepts and examine the literature to such a detailed extent, particularly when it seems likely that the vast majority of publications are produced in good faith and come from institutions of higher learning. Ioannidis,<sup>40</sup> however, has argued that most published research findings are false. Ioannidis' estimate is sensitive to the pretest probability of the hypothesis being true, and a low estimate of this value will lead to a higher proportion of papers' conclusions being classed as false. Nevertheless, as will be discussed later, current research into reproducibility of findings has provided more direct evidence indicating a significant proportion of findings are false in that they cannot be reproduced. In Ioannidis' common sense view, a research finding is less likely to be true when effect sizes are small, when there are a large number of tested hypotheses that have not been preselected, and when there are great flexibilities in designs, definitions, outcomes, and data analyses. Other problems impacting the truth of the conclusion are financial issues and other interests and prejudices as well as the number of teams in a field chasing statistical significance. I believe it is unlikely that most research findings are false, because if they were there would be more papers reporting failure to confirm results (though admittedly publishing such negative results can be difficult) and many fewer papers confirming—albeit often indirectly—replication. Nevertheless, the considerations listed by Ioannidis serve to warn readers of the dental

and medical literature that there is a shortage of well-documented threats to truth.



### Is it important?

The importance of a paper cannot be tested in a completely objective manner. Maxwell<sup>41</sup> has argued—in my opinion, persuasively—that real progress in science is assessed in terms of the amount of valuable factual truth that is being discovered and that the accumulation of vast amounts of trivia (even if factually correct) does not amount to progress. The problem is that value judgments are highly subjective. One approach to measuring impact of a paper is the number of citations to the paper, an aspect that will be discussed in chapters 5 and 22. Many scientists have accepted this criterion and include the citation record of their papers in their curriculum vitae or include indices derived from their citation record such as the h-index (discussed in chapter 5).

One can speculate about what qualities an ideal evaluator should have. Beveridge<sup>42</sup> has suggested the concept of scientific taste, which he described as a sense of beauty or esthetic sensibility. Beveridge explained scientific taste by stating that:

The person who possesses the flair for choosing profitable lines of investigation is able to see further where the work is leading than are other people because he has the habit of using his imagination to look far ahead, instead of restricting his thinking to established knowledge and the immediate problem.

A person with scientific taste would be a good judge of the importance of a scientific paper. Traditionally, the skill of judgment is learned in the apprentice-master relationship formed between graduate student and supervisor. Techniques may come and go, but judging what is important and how it can be innovatively studied are the core business of scientists, and these skills are learned similarly to a child learning his prayers at his mother's knee: Graduate students hone their critical skills in the supervisor's office or lab meetings. Thus, much importance is attached to the pedigree of a scientist, and some scientists take pride in tracing their scientific pedigrees to leading figures in a field of study.

Given the large variation in laboratory and supervisor quality, there will always be significant differences in judgment. This diversity is evident in an extensive

study of proposals submitted to the National Science Foundation. The study found that getting a research grant significantly depends on chance, because there is substantial disagreement among eligible reviewers, and the success of a proposal rests on which reviewers happen to be accepted.<sup>43</sup> Moreover, there is evidence that complete disagreement between pairs of referees assessing the same paper is common.<sup>43</sup> In biomedical science, the frequency of agreement between referees was not much better than that which would be expected by chance.<sup>44</sup> Hence, it appears that objective and absolute criteria for the evaluation of a paper prior to publication are not available. Chapter 22 attempts to cultivate the skill of judgment by providing information on recognized sources of errors in judgments as well as citation analysis, a technique that can be used to access broadly based scientific assessments of published works.

## References

1. Chargaff E. Triviality in science: A brief meditation on fashions. *Perspect Biol Med* 1976;19:324.
2. Norris SP. Synthesis of research on critical thinking. *Educ Leadersh* 1985;42:40–45.
3. Kurland DJ. I Know What It Says—What Does It Mean? *Critical Skills for Critical Reading*. Belmont, CA: Wadsworth, 1995:164.
4. Butler CB. New ways of seeing the world. In: Butler CB (ed). *Post-modernism: A Very Short Introduction*. New York: Oxford University, 2002:37–42.
5. Bronowski J. The common sense of science. In: Bronowski J (ed). *The Common Sense of Science*. New York: Vintage, 1967:97–118.
6. Chambers DW. Lessons from badly behaved technology transfer to a professional context. *Int J Technol Transfer Commercialisation* 2005;4:63.
7. Chambers DW. Changing dental disease patterns. *Contact Point* 1985;63:1–17.
8. Fish W. Framing and testing hypotheses. In: Cohen B, Kramer IRH (eds). *Scientific Foundations of Dentistry*. London: William Heinemann, 1976:669.
9. Perkins DN, Allen R, Hafner J. Difficulties in everyday reasoning. In: Maxwell W (ed). *Thinking: The Expanding Frontier*. Philadelphia: Franklin Institute, 1983:177.
10. Zipf GK. *Human Behaviour and the Principle of Least Effort*. Cambridge: Addison-Wesley, 1949.
11. Pratkanis AR, Aronson E. *Age of Propaganda: The Everyday Use and Abuse of Persuasion*. New York: WH Freeman, 2001:38.
12. Kahneman D. Attention and effort. In: Kahneman D (ed). *Thinking, Fast and Slow*. Toronto: Anchor Canada, 2013:31–38.
13. Risen J, Gilovich T. Informal logical fallacies. In: Sternberg RJ, Roediger HL, Halpern DF (eds). *Critical Thinking in Psychology*. Cambridge: Cambridge University, 2007:110–130.
14. Garfield E. Current comments: Is the ratio between number of citations and publications cited a true constant? *Curr Contents* 1976;6:5–7.
15. Glenny A, Hooper L. Why are systematic reviews useful? In: Clarkson J, Harrison JE, Ismail A (eds). *Evidence Based Dentistry for Effective Practice*. London: Martin Dunitz, 2003:55.
16. Garfield E. The literature of dental science vs the literature used by dental researchers. In: Garfield E (eds). *Essays of an Information Scientist*. Philadelphia: ISI, 1982:373.
17. Relman AS. *Journals*. In: Warren KS (ed). *Coping with the Biomedical Literature*. New York: Praeger, 1981:67.
18. Worthington H, Clarkson J. Systematic reviews in dentistry: The role of the Cochrane oral health group. In: Clarkson J, Harrison JE, Ismail A (eds). *Evidence Based Dentistry for Effective Practice*. London: Martin Dunitz, 2003:97.
19. Chambers DW. Habits of the reflective practitioner. *Contact Point* 1999;79:8–10.
20. Day RA. *How to Write and Publish a Scientific Paper*. Philadelphia: ISI, 1979:2.
21. DeBakey L. *The Scientific Journal. Editorial Policies and Practices: Guidelines for Editors, Reviewers, and Authors*. St Louis: Mosby, 1976:1–3.
22. Simonton DK. *Creativity in Science: Chance, Logic, Genius, and Zeitgeist*. Cambridge: Cambridge University, 2004:85–86.
23. Goodman SN, Berlin J, Fletcher SW, Fletcher RH. Manuscript quality before and after peer review and editing at *Annals of Internal Medicine*. *Ann Intern Med* 1994;121(1):11.
24. Giannobile WV. Editor's Report for the *Journal of Dental Research*—2015. <http://dentalresearchblog.org/jdr/?p=338>. Accessed 18 July 2018.
25. Tacker MM. Parts of the research report: The title. *Int J Prosthodont* 1990;3:396–397.
26. Chernin E. Producing biomedical information: First, do no harm. In: Warren KS (ed). *Coping with the Biomedical Literature*. New York: Praeger, 1981:40–65.
27. Day RA. *How to Write and Publish a Scientific Paper*. Philadelphia: ISI, 1979:72.
28. Cialdini RB. The science of persuasion. *Sci Am* 2001;284:76–81.
29. Gross AG. *The Rhetoric of Science*. Cambridge: Harvard University, 1990.
30. Gopen G, Swan J. The science of scientific writing. *Am Sci* 1990;78:550–558.
31. Strunk W, White EB. *The Elements of Style*, ed 4. Boston: Allyn and Bacon, 1999.
32. Zinnser W. *On Writing Well*, ed 2. New York: Harper and Row, 1980.
33. Walton D. Argument as reasoned dialogue. In: Walton D (ed). *Informal Logic: A Pragmatic Approach*, ed 2. New York: Cambridge University, 2008:1–34.
34. Toulmin SE. *The Uses of Argument*, updated ed. Cambridge: Cambridge University, 2003.
35. Lang TA. *How to Write Publish and Present in the Health Sciences: A Guide for Clinicians and Laboratory Researchers*. Philadelphia: American College of Physicians, 2010:41.
36. Ziman J. *Reliable Knowledge: An Exploration of the Grounds for Belief in Science*. Cambridge: Cambridge University, 1978:7.
37. DeBakey L. *The Scientific Journal: Editorial Policies and Practices: Guidelines for Editors, Reviewers, and Authors*. St Louis: Mosby, 1976:1.
38. Kuhn TS. *The Structure of Scientific Revolutions*, ed 2. Chicago: Chicago University, 1970:184.
39. Fisher RA. *The Design of Experiments*, ed 8. Edinburgh: Oliver & Boyd, 1953:184.
40. Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2:e124.



## REASONS FOR STUDYING CRITICAL THINKING

41. Maxwell N. Articulating the aims of science. *Nature* 1977;265:2.
42. Beveridge WJ. *The Art of Scientific Investigation*. New York: Vintage, 1950:106.
43. Cole S, Cole JR, Simon GA. Chance and chaos in peer review. *Science* 1981;214:881.
44. Gordon M. Evaluating the evaluators. *New Sci* 1977;61:10.



Page numbers followed by “f” denote figures; those followed by “t” tables; and those followed by “b” denote boxes

## A

Abduction, 76, 84–85  
 Aberrant behavior, 17–18  
 Aberration of logic, 85  
 Absorbance errors, 289  
 Abstract  
   function of, 34–35  
   structured, 35–36  
 Academic science, 23, 315, 328  
 Acceptance date, for scientific papers, 34  
 Accountability, 28  
 Accuracy  
   of measurements, 155, 201b  
   of ratios, 161  
   of scientific writing, 316  
   precision versus, 164, 164f  
 Active control, 278–279  
 Active reading, 32–33  
 Ad hominem, 122  
 Ad populum, 123  
 Adaptive sampling, 291  
 Advanced Google search, 72, 72f  
 Affective error, 320  
 AGREE II appraisal tool, 59  
 Agreement indices, 210  
 Alternative health industry, 119  
 Alternative hypothesis, 132  
 Altmetrics, 65  
 AMSTAR, 58  
 Analogy, 87–88  
 Analysis of variance. *See* ANOVA.  
 Analytic variation, 154  
 AND operator, 68  
 Anecdotal evidence, 110  
 ANOVA  
   analysis of errors by, 171–172  
   factorial design analysis, 296  
   Friedman two-way, 158t  
   one-way, 159t  
   reliability estimations, 212  
   two-factor, 159t  
   units of measurement, 160  
 Antecedent, 91  
 Applied kinesiology, 121  
 Applied research, 280  
 Argument(s)

  by analogy, 76  
   from authority, 89–91  
   chains, 82  
   deductive, 77  
   definition of, 74  
   formal analysis of, 83  
   inductive. *See* Inductive arguments.  
   standards for assessing, 74  
   tu quoque, 122  
 Argumentation maps, 316–318, 317f  
 Aristotle, 40–42  
 Arrows, in illustrations, 187  
 Articulation, 46  
 Artifact, 165, 181  
 Assertability question, 35, 76  
 Assertions, 75  
 Assessment, 198  
 Assimilation bias, 320  
 Association, 257  
 Assumptions, 271–272  
 Audience, 45–46, 50–51  
 Auriculotherapy, 121  
 Author(s)  
   goal for, 4  
   institutional affiliations of, 34  
   of scientific papers, 33–34  
   ordering of, 34  
   referees versus, 7–8  
   significance of, 34  
 Authority, 53, 89–91  
 Author-level citation-based metrics, 65  
 Auxiliary hypotheses, 92  
 Availability bias, 338  
 Availability heuristic, 321  
 Axelrod, Robert, 16

## B

Babbage, Charles, 17  
 Bandwagons, 52  
 Bar graphs, 191–192  
 Bayesian approach, 312–313  
 Bean machine, 168  
 Begging the question, 124  
 Benevolent behaviors, 25–26  
 Benford’s law, 310  
 Bernard, Claude, 111

## Bias

  assimilation, 320  
   availability, 338  
   commission, 320  
   confirmation, 246–248, 320  
   definition of, 265  
   detection, 270, 306  
   information, 267  
   observational error, 307  
   reporting in, 118  
   response, 38, 307  
   results affected by, 166  
   sampling, 307  
   selection, 269, 306  
   treatment allocation, 307  
   unintentional, 18–19  
   volunteer, 306  
   withdrawal, 306  
 Bimodal distribution, 161  
 Binomial coefficient, 134  
 Binomial distribution, 133, 150  
 Binomial effect size display, 181  
 Biologic variation, 154, 208  
 BioMed Central, 62  
 Bivalent experiments, 292  
 Bivariate frequency distribution  
   tables, 142. *See also*  
   Contingency tables.  
 Black-box approach, 89  
 Blinding, 284–285  
 Blocking, 293  
 Blunders, 167  
 Bohr model, 87–88  
 Books, 60  
 Boolean operators, 67–68, 69f  
 Burt, Cyril, 16–17

## C

CADTH. *See* Canadian Agency for  
 Drugs and Technologies in  
 Health.  
 Calibration errors, 170–171  
 Canadian Agency for Drugs and  
 Technologies in Health, 62, 62b  
 Canadian Institutes of Health  
 Research, 63  
 Canons of induction, 99

Carry-across effect, 298  
 Cartesian skepticism, 76  
 Case study, 250, 254–255  
 Case-control designs, 266–268, 267f  
 Case-control study, 237  
 Case-series analysis, 255  
 Categorical statements, 79–80  
 Categorical syllogisms, 80–82  
 Causal chain, 102f, 103  
 Causal conclusions, 244  
 Causal factors, 100  
 Causal reasoning, 321  
 Causality, 98  
 Causation  
   assessment of, 100–101, 106–107  
   convincing variables and, 101–106  
   correlation and, 270  
   criteria for, 105–106  
   direct, 102  
   indirect, 102  
   patterns of, 102–105  
   pleiotropy, 102–103  
   problems in determining, 270–272  
   Sackett's diagnostic tests for, 105–106  
 Cause-and-effect hypotheses, 97–100  
 Cause-and-effect relationships, 104f  
   convincing evidence for  
     establishing, 106  
     ends and means, 107–108  
     interpretation of, 261  
     investigation of, 105–106  
 Cause-effect correlations, 125, 261  
 Central limit theorem, 168  
 Central persuasion, 47  
 Central tendency measures, 184–185  
 Chains, 82, 103  
 Chartjunk, 188  
 Cheating, 18  
 Chiropractic treatment, 113  
 Chi-squared test, 158t  
 Chronologic approach, 44  
 Cialdini principles of influence, 51–53  
 Cicero, 43  
 Circular reasoning, 124  
 Citation analysis, 10, 24, 34, 63–65, 326–329  
 Citation frequency, 329  
 Citation tracking, 71  
 Clarity, of scientific writing, 316  
 Classical probability, 129  
 Cleveland's hierarchy of graphical perception, 190–192  
 Clinical decision making, 331–333.  
   *See also* Decision making.  
 Clinical dentistry fallacies, 124–125  
 Clinical diagnosis–decision algorithm, 203f  
 Clinical judgment, 315, 322–323  
 Clinical outcomes, 338  
 Clinical practice guidelines, 57  
 Clinical problem, 336

Clinical studies, 22  
 Cochrane Collaboration, 58  
 Cochrane Database of Systematic Reviews, 59  
 Cochrane Library, 3  
 Coefficient of determination, 262  
 Coefficient of variation, 208–209  
 Cognitive dissonance, 48–49  
 Cognitive miser, 2  
 Cognitive response, 49–50  
 Cohen kappa, 259–260  
 Cohort, 268  
 Cohort studies, 237, 268–269  
 Cold fusion, 19  
 Collaboration, 16  
 Collective judgment, 328  
 Commission bias, 320  
 Common sense, 1, 13, 98  
 Common theory-evidence relationships, 112–114  
 Communalism, 14  
 Comparisons, 176  
 Complementary and alternative medicine, 115–116  
 Complex etiology, 104–105  
 Compliance of subjects, 285  
 Composition fallacy, 122–123  
 Concept mapping, 316  
 Conceptual model, 88  
 Concurrent validity, 156  
 Conditional probability, 130, 205–206  
 Conditional statements, 91–94  
 Confidence intervals, 147–150, 169  
 Confidence limits, 148, 168  
 Confirmability, 251  
 Confirmation bias, 246–248, 320  
 Confounding, 270–271  
 Confounding variable, 101, 260, 263  
 Conjunctions, 102f, 103  
 Consequences-based validity, 156  
 Consequent, 91  
 Consistency, 51–52  
 CONSORT guidelines, 36  
 Construct-based validity, 156, 244  
 Content-based validity, 156  
 Contiguity, 99  
 Contingency tables, 142–143, 156f  
 Continuous response variables, 284  
 Continuous variables, 145  
 Contrast, 51  
 Contributory causes, 100–101  
 Controls, 277–279, 293  
 Controversial care, 121  
 Convenience sample, 242  
 Conventionalism, 29, 95  
 Convincing evidence, 106  
 Cooperation, 16  
 Correlation  
   causation and, 270  
   description of, 257, 260–261  
   regression versus, 261–264  
 Correlation coefficient, 155, 172, 209–210, 261–262

Correlational cross-sectional study, 258, 258t  
 Cost-effectiveness analysis, 348f  
 Counterbalanced designs, 297–298  
 Covariates, 263  
 CPGs. *See* Clinical practice guidelines.  
 Cranial osteopathy, 121  
 Cranks, 114–115  
 Credibility, 25, 43, 47, 90, 250  
 Critical theory, 251  
 Critical thinking  
   decision making and, 331–332  
   description of, 1, 32, 315–316  
 Crossover design, 297t, 297–298  
 Crossoveritis, 298  
 Cross-sectional designs, 238  
 Cross-sectional survey, 258t, 258–265  
 Cross-validation, 88–89  
 Cut-off lines, 181

## D

Darwin, Charles, 332  
 Data. *See also* Results.  
   analysis of, 182–183, 310  
   derived measures for, 180–181  
   from experiments, 285–286  
   hiding of, 193f, 194  
   magnifying a small amount of, 194, 195f  
   manipulation of, 180  
   normalization of, 180  
   presentation of, 178–180  
   qualitative, 251  
   quality of, 175–176  
   unidimensional, 195–196  
 Data density index, 190, 194  
 Data reporting  
   conventions in, 186  
   minimum requirements for, 183  
   significant digits, 186  
 Databases, 66  
 Data-ink ratio, 190  
 Decision analysis, 201b, 202f  
 Decision balance sheet, 336  
 Decision making  
   critical thinking and, 331–332  
   decision tree analysis for, 347, 350–351  
   description of, 86  
   evidence-based, 335, 350–351  
   medical, 335  
   one-reason, 332, 351  
   in patient care, 333–350  
 Decision tree, 336–337  
 Decision tree analysis  
   decision-making uses of, 350–351  
   definition of, 335  
   economic analysis and, 350  
   example of, 343  
   limitations of, 347–350  
   steps involved in, 335–345  
 Deduction, 42



Deductive argument, 77  
 Deductive logic  
   description of, 76–78  
   errors in, 82–83  
   in scientific articles, 77–78  
   syllogisms. *See* Syllogisms.  
 Degrees of freedom, 149, 185, 303–304  
 Dentistry, 1–4  
 Dependability, 251  
 Dependent variables, 274–275  
 Descriptive statistics, 128, 305  
 Descriptive study, 237  
 Detection bias, 270, 306  
 Determinate error, 164, 165–167  
 Developmental research, 280  
 Diagnosis  
   algorithm for, 203f  
   decision analysis for, 201b  
   principle of, 200  
   process of, 201–204  
 Diagnostic rates, 207t  
 Diagnostic tests and measurements  
   decision analysis, 201b  
   description of, 200  
   reliability of, 206–214  
   uses of, 204  
 Diagnostic tool, 204–206  
 Diagrams, 187  
 Dichotomous response variable, 284  
 Digital photography, 174  
 Digression, 50  
 Direct causes, 102  
 Discrete variables, 130, 145  
 Discussion/conclusions section, of  
   scientific papers, 38  
 Disease  
   description of, 200  
   pretest probability of, 202  
 Disinterestedness, 14–15  
 Disjunctive syllogism, 82, 124, 135  
 Dispersion measures, 185–186  
 Dispositio, 44  
 Distribution  
   a priori, 131  
   bimodal, 161  
   binomial, 133, 150  
   frequency, 131, 134  
   Gaussian, 141, 146, 157, 167  
   normal. *See* Normal distribution.  
   Poisson, 140–142  
   probability, 133–134  
   reference, 133  
   Student, 149  
   t, 149  
   uniform, 145  
 Doctrinaire rules-based approach,  
   111  
 Double-blind treatment/trial, 284,  
   291  
 DTA. *See* Decision tree analysis.

## E

EBD. *See* Evidence-based dentistry.

Ecologic study, 265–266, 266f  
 Economic analysis, 345–347, 350  
 Editorial consultants, 4  
 Editors, 5–6  
 Effect-size approach, 304–305  
 Einstein, Albert, 17, 93  
 Elementary outcomes, 129  
 Emergent design, 250  
 Emotional appeal, 42  
 Endogenous research, 251  
 Ends and means, 107–108  
 Enthymeme, 42  
 Epidemiologic measures, 245. *See also*  
   Prevalence.  
 Epistemology, 2  
 Equivocation, fallacy of, 45, 316  
 Error(s)  
   analytic variation, 154  
   ANOVA used for analysis of,  
     171–172  
   assignable causes of, 167  
   biologic variation, 154  
   blunders, 167  
   calibration, 170–171  
   causes of, 165  
   data analysis, 310  
   in deductive logic, 82–83  
   determinate, 164, 165–167  
   distributing of, 289  
   indeterminate, 164, 167–169  
   law of propagation of, 169–170  
   management of, 288–291  
   measurement of, 289–290  
   method of least squares, 168f, 169  
   normal law of, 167–168  
   observation-related, 165  
   of combined measurements,  
     169–170  
   personal, 166  
   random, 164, 167, 279  
   reducing the size of, 290–291  
   risk of, 133  
   standard. *See* Standard error.  
   systematic, 164, 165–166, 279  
   type I, 138, 138t  
   type II, 137, 138t  
   units as cause of, 310  
 Error sum of squares, 169  
 Estimation, 149–150  
 Ethical appeal, 42–44  
 Ethnography, 251  
 Ethos, 42–44  
 Event, 129  
 Evidence  
   anecdotal, 110  
   common theory and, relationships  
     between, 112–114  
   comprehensiveness of, 116  
   convincing, 106  
   fallacies of, 75–76  
   incomplete, 86t  
   insufficient or inappropriate, 75–76  
   level of, 3t  
   line of, 23  
 Evidence pyramid, 56, 57f

Evidence-based decision making,  
   335, 350–351  
 Evidence-based dentistry,  
   110–111  
 Evidence-based medicine, 110–111  
 Examiner variability, 209–214  
 Existential assumption, 81  
 Expected utility value, 341–344, 349  
 Experiment(s). *See also* Research.  
   bivalent, 292  
   controls, 277–279  
   data quality, 285–286  
   dependent variables, 274–275  
   functional, 292  
   independent variables, 274–275  
   measurement, 279–280  
   random, 129  
   requirements for, 276–280  
   research, 280–281  
   tactics, 281–286  
   validity of, 280  
   variables, 286  
 Experiment design  
   counterbalanced, 297–298  
   crossover, 297t, 297–298  
   description of, 288  
   error management, 288–291  
   factorial, 295–296  
   matched-group, 293  
   pre-experimental, 298–299  
   quasi-experimental, 298–299  
   randomized controlled trials, 235–  
     236, 291–292, 299–300  
   split-mouth, 298  
   true, 298–299  
 Experts, 89–90  
 Explanation, 108  
 Explanatory data analysis, 182  
 Explanatory trials, 238  
 Exploratory data analysis, 182–183  
 Exploratory research, 22, 280  
 Exposure measurement error,  
   264–265  
 External validity, 242–244, 298  
 Extrapolation, 292

## F

Face validity, 156  
 Facet approach, 44  
 Factorial designs, 295–296  
 Failure fallacy, 125  
 Fallacies  
   genetic, 123  
   of affirming the consequent, 91  
   of apples and oranges, 123–124  
   of biased statistics, 305–307  
   of clinical dentistry, 124–125  
   of composition, 122–123  
   of concomitant variation, 260  
   of equivocation, 45, 316  
   of evidence search and  
     interpretation, 320  
   of existential assumption, 81  
   of failure, 125  
   of informal logic, 121–125

of insufficient or inappropriate evidence, 75–76  
of insufficient statistics, 302–303  
of no evidence, 75  
of origin, 122  
of special pleading, 122  
of success, 124–125  
UFO, 124, 307  
Fast and frugal heuristics, 331–332  
FFH. *See* Fast and frugal heuristics.  
Figures. *See* Illustrations.  
File drawer problem, 309  
Fisher exact test, 158t  
Focal infection theory/hypothesis, 2, 321  
Follow-up (cohort) design, 268–269  
Formal deductive arguments, 77  
Franklin, Benjamin, 333, 350  
Fraud, 17–18, 37  
Freidman two-way ANOVA, 158t  
Frequency distributions, 131, 134. *See* also Contingency tables.  
Frequentist probability, 129  
Frequentists, 338  
Functional experiments, 292

## G

Galton board, 168, 168f  
Gaussian distribution, 141, 146, 157, 167  
Generality, 46–47  
Generalizability, 252  
Genetic fallacy, 123  
Goodness of fit, 138–142  
Google Scholar, 70–72, 71f  
Graphs  
bar, 191–192  
Cleveland's hierarchy of graphical perception, 190–192  
deception in, 192–196  
description of, 187  
exploratory data analysis use of, 183  
layering of, 189  
parallelism of, 189  
presentation techniques for, 187–188  
separation of, 189  
strategies for using, 188–189  
Tuft's evaluative ratios for, 189–190  
zero value of the ordinate, 194, 195f  
Green open access, 63  
Grey literature, 60, 62, 62b, 66, 86  
Grounded theory, 250–251  
Groupthink, 52  
Guidelines, 59, 59b

## H

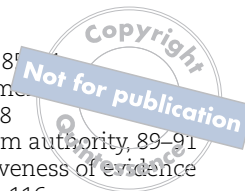
Health-state utility, 340–341

Heterogeneity of treatment effects, 237, 300  
Heuristics, 50–51, 251, 321–322, 325, 331–332  
High-impact research, 15–16  
H-index, 64–65  
Histogram, 192. *See* also Relative frequency histogram.  
Historical-controlled trials, 237  
Homeopathy, 113  
Hume, David, 98–99, 261  
Humility, 15  
Hypercompetition, 27–28  
Hypothesis  
acceptance of, 24  
alternative, 132, 307  
auxiliary, 92  
cause and effect, 97–100  
comprehensiveness of, 93  
criticizing of, 94  
definition of, 91  
description of, 23  
disproving of, 91–92  
evaluation of, 93  
internal consistency of, 93  
null, 132, 135, 308–309  
objections to, 94b  
of no effect, 312  
simplicity of, 93  
successful predictions, 93  
testability of, 93, 116  
Hypothesis testing  
confidence intervals used to predict results of, 149  
description of, 88–89, 91  
steps involved in, 131–137  
Hypothetical inference, 91–94

## I

ICER. *See* Incremental cost-effectiveness ratio.  
Idealism, 21  
Illustrations  
arrows in, 187  
deception in, 192–196  
description of, 186–187  
graphs. *See* Graphs.  
presentation techniques for, 187–188  
Impact factor, for journals, 64  
Incidence, 245  
Incomplete evidence, 86t  
Incremental cost-effectiveness ratio, 346, 346t  
Independent events, 130  
Independent variables, 274–275  
Indeterminacy principle, 165  
Indeterminate errors, 164, 167–169  
Indirect causes, 104  
Induction  
Bayesian approach to, 312–313  
canons of, 99  
description of, 42  
by enumeration, 302

principles of, 85  
Inductive arguments  
analogy, 87–88  
argument from authority, 89–91  
comprehensiveness of evidence regarding, 116  
criterion for acceptability of, 116  
logic of criticism of, 95–96  
models, 88–89  
statistical inference as, 302–305  
Inductive logic  
additional evidence in, 85–86  
arguments. *See* Inductive arguments.  
description of, 76, 85  
Inferential statistics, 128, 305  
Infinite dilution, 19  
Informal logic, 77  
Information bias, 267  
Information resources  
citation analysis, 10, 24, 34, 63–65, 326–329  
description of, 3–4  
Information searches  
Advanced Google, 72, 72f  
Boolean operators, 67–68, 69f  
conducting of, 65–72  
databases, 66  
defining of question for, 65–66  
Google Scholar, 70–72, 71f  
help in, 55–56  
keyword, 66–67, 69f  
parentheses in, 69  
phrases in, 69  
PubMed, 69–70, 70f  
resources used in, 66  
subject headings, 66–67, 69  
techniques for, 66–69  
tips for, 72–73  
truncation, 68–69  
Information sources  
books, 60  
filtered, 56  
grey literature, 60, 62, 62b, 66, 86  
guidelines, 59, 59b  
journals, 60  
knowledge synthesis reviews, 57–59, 59b  
open access literature, 62–63  
overview of, 56–57  
peer review, 60  
pre-appraised clinical tools, 59–60  
selection of, 66  
systematic reviews, 57–59  
Inhibitors, 283  
Insufficient but redundant part of an unnecessary but sufficient condition, 100  
Insufficient or inappropriate evidence, 75–76  
Integrity, 37–38, 173–175  
Intention-to-treat analysis, 285  
Interaction effect, 296  
Interactions, 102f, 103  
Interestingness, 47



Inter-examiner comparison, 209  
Internal validity  
  description of, 239, 244, 298  
  threats to, 239–241  
International Committee of Medical  
  Journal Editors, 33  
Interpolation, 292  
Interquartile range, 185  
Interval scale, 159, 159t  
Intra-examiner comparison, 209  
Introduction section, of scientific  
  papers, 36  
Inventio, 41–44  
Irreproducibility epidemic, 26

## J

Joint stretching, 88  
Journals  
  as information sources, 60  
  citation analysis of, 64  
  description of, 2, 4  
  impact factor for, 64  
  open access publishing models  
  for, 62  
  refereed, 60, 61b  
  reporting practices of, 27  
Judgments  
  balanced, 318–319  
  clinical, 315, 322–323  
  scientific, 315  
  under uncertainty, 319–323  
Judicial precedent, 87  
Jujitsu method, 282

## K

Kappa, 210–211  
Keyword searches, 66–67, 69f  
Knowledge synthesis reviews, 57–59,  
  59b  
Koch's postulates, 105  
Kruskall-Willis test, 158t

## L

Lack of objectivity, 116–117  
Law, analogy and, 87  
Law of cognitive response, 49–50  
Law of propagation of errors, 169–  
  170  
Law of small numbers. *See* Poisson  
  distribution.  
Letter to the editor, 24  
Level of evidence, 3t  
Level of significance, 133  
Librarians, 55  
Lie factor, 46, 190, 195  
Life history, 251  
Liking, 52–53  
Logarithmic scales, 193  
Logic  
  abductive, 84–85  
  aberration of, 85  
  abuses of, 121–125

deductive. *See* Deductive logic.  
inductive, 76, 85–91  
informal, 77  
reasons for studying, 84  
in scientific writing, 316  
semi-formal, 77  
Logistic regression, 264  
Logos, 42  
Longitudinal study, 240

## M

Magnitude, 46  
Make-sense epistemology, 2  
Malevolent behaviors, 25–26  
Managing editors, 6  
Mann-Whitney U test, 158t  
Mapping, 176–178  
Markov chain Monte Carlo, 313  
Matched-group experiment designs,  
  293  
Matching, 293  
Materials and methods section,  
  36–37  
McLuhan, Marshall, 188  
MCMC. *See* Markov chain Monte  
  Carlo.  
McNemar chi-squared test, 158t  
Mead, Margaret, 19  
Mean, 184, 184f  
Means and ends, 107–108  
Measurements  
  accuracy of, 155  
  aggregated, 181–183  
  choice of, 175  
  combined, 169–170  
  definition of, 152  
  diagnostic. *See* Diagnostic tests  
  and measurements.  
  errors of. *See* Error(s).  
  operational definitions, 152–153  
  precision of, 154, 164, 164f, 209  
  reliability of, 172, 206–214  
  resolution, 153–154  
  results of. *See* Results.  
  scales for comparisons among. *See*  
  Scales.  
  true score theory of, 154  
  units of, 160–161  
  validity of, 155–157  
Measures of central tendency,  
  184–185  
Measures of dispersion, 185–186  
Mechanism, 88  
Median, 184, 184f  
Medical decision making, 335  
MEDLINE, 67  
Mendel, Gregor, 138–140  
Mental processes, 3  
Mere, Chevalier de, 129  
MeSH, 35, 67  
Meta-analysis, 57, 149, 323, 325–326  
Method of difference, 277  
Method of least squares, 168f, 169

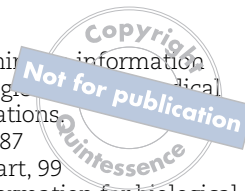
MIBBI. *See* Minimum information  
  for biological and biomedical  
  investigations.  
Micrograph, 187  
Mill, John Stuart, 99  
Minimum information for biological  
  and biomedical investigations,  
  37  
Misclassification, 264–265  
Mixed hypothetical syllogism, 81  
Mode, 184, 184f  
Models, 89, 320–321  
Modus tollens, 92  
Moral algebra, 333, 350  
Multidisciplinary collaborative  
  research, 42  
Multiple linear regression, 263  
Mutually exclusive events, 129

## N

Narrative reviews, 58, 58b  
Naturalistic inquiry, 250  
Necessary condition, 99–100  
Negative control, 278  
Negative predictive value, 201b, 207t  
Negative results papers, 95  
Neyman-Pearson approach, 311  
Nominal scale, 157  
Nondisjoint outcomes, 129  
Nonparametric statistics, 140  
Nonparametric tests, 157, 158t  
Normal distribution  
  advantages of, 150  
  definition of, 146  
  sampling from, 148–149  
  shape of, 184  
  standard, 147  
Normal law of error, 167–169  
Normalization, 180  
Nosographic rates, 207t  
NOT operator, 68  
Novel indices, 181  
Null hypothesis, 132, 135, 308–309  
Numerical method, 110  
Numerical quality, 176  
Nutrition quackery, 121

## O

Objective information, 198  
Objectivity, 7, 16–19, 116–117, 247  
Obsequence bias, 307  
Observation(s)  
  authority versus, 89  
  case study design affected by,  
  254–255  
  classification of, 22, 248  
  discoveries through, 245  
  error caused by, 165  
  limitations of, 21  
  selection of, 21–22, 249  
  unreliability of, 307  
Observational error bias, 307  
Observational reactivity, 249



Observational studies, 237, 271  
 Observation-description strategy  
   advantages of, 248–249  
   case study, 254–255  
   case-series analysis, 255  
   in clinical interventions, 253–255  
   epidemiologic measures, 245  
   operational considerations for, 248  
   technologic influences, 249  
   weaknesses associated with, 249  
 Occam's razor, 93, 135  
 Odds ratio, 259  
 One-reason decision making, 332, 351  
 One-tailed tests, 137, 311  
 One-way ANOVA, 159t  
 One-way sensitivity analysis, 344  
 Online profiles, 65  
 Open access, 62–63  
 Open data, 62–63  
 Open science, 62–63  
 Open-blind study, 284  
 Open-mindedness, 74  
 Openness, 25–29  
 Operational definitions, 152–153  
 OR operator, 67–68  
 ORCID, 65  
 Ordinal scale, 157–159, 158t  
 Organizational schemes, 44  
 Organized skepticism, 14, 24  
 Origin fallacy, 122  
 Originality, 9, 15, 280  
 Outcome(s)  
   in dentistry, 339t  
   description of, 129–130  
   patient-centered, 339–340  
   probability of, 338  
 Outcome measures, 284  
 Outcome of interest, 105  
 Outcome variables, 22, 263  
 Overtreatment, 125

## P

P value, 135  
 Paired t test, 159t  
 Pairing, 293  
 Papers. *See* Scientific papers.  
 Parallelism, 189  
 Parameter, 140, 274  
 Parametric tests, 157  
 Parent population, 305  
 Parentheses, 69  
 Pareto's law, 167  
 Participatory action research, 251  
 Pascal, Blaise, 129  
 Passive voice, 16, 43  
 Pathologic science, 19  
 Pathophysiologic approach, 110–111  
 Pathos, 42  
 Patient-centered outcomes, 339–340  
 Pearson correlation coefficient, 209  
 Peer review, 60  
 Peer-review ratings, 5  
 Percent agreement, 210

Percent reproducibility, 212  
 Peripheral persuasion, 47  
 Personal accountability, 28  
 Personal errors, 166  
 Personal experience, 117  
 Perspective mania, 196  
 "Persuading with pap," 187–188  
 Persuasion, 46–49, 173  
 Phenomenology, 251  
 Physical model, 88  
 PICO, 56, 65–67, 336, 339  
 Placebo effect, 117–118, 124  
 Plato, 41  
 Pleiotropy, 102f, 102–103  
 Point-of-care tools, 59–60  
 Poisson distribution, 140–142  
 Population  
   definition of, 184, 305  
   of known dispersion, 147–148  
   target, 305, 307  
   variability in, 185  
   whose dispersion is not known "a priori," 149  
 Population overlap, 308  
 Positive control, 278–279  
 Positive predictive value, 207t  
 Positivism, 29, 95  
 Post hoc, 123  
 Posttest likelihoods, 201b  
 Posttest probability, 207t  
 Pragmatic model, of reasoning, 8  
 Pragmatic trials, 238  
 Pre-appraised clinical tools, 59–60  
 Precision, 154, 164, 164f, 209, 279  
 Predictive validity, 155  
 Pre-experimental designs, 298–299  
 Premises, 74, 82  
 Pretest probability, 201–202, 207t  
 Prevalence, 201b, 205, 245  
 Price index, 326  
 Principal principle, 48  
 Principle of Least Effort, 2  
 Principle of social proof, 52  
 PRISMA, 58, 59b  
 Probabilistic cause, 100  
 Probability  
   calculations, 129–130  
   classical, 129  
   conditional, 130, 205–206  
   conditions associated with, 205  
   definition of, 129  
   frequentist, 129  
   history of, 129  
   measuring, 205  
   subjective, 129  
 Probability density, 145  
 Probability distribution, 133–134  
 Problem solving, 88  
 Profiles, online, 65  
 Proof by exclusion, 124  
 Proof of elimination, 124  
 Propagation of errors theorem, 169–170  
 Proportion agreement, 210  
 Prospective cohort study, 238

Psychologic studies, 237  
 Ptolemy, Claudius  
 Public Library of Science, 62  
 Publication  
   authors, 4, 7–8  
   description of, 23  
   editors, 5–6  
   nonrefereed journals, 5  
   readers of, 8  
   refereed journals, 5  
   referees, 5–8  
 Publishing pressure, 17, 27  
 PubMed, 69–70, 70f  
 PubMed Central, 62  
 Pure hypothetical syllogism, 81  
 Pure research, 280  
 Purposive sampling, 250

## Q

QALY. *See* Quality adjusted life years.  
 QAPY. *See* Quality adjusted prosthetic years.  
 Quackery, 115, 120–121  
 Quacks  
   complementary and alternative medicine, 115–116  
   definition of, 114–115  
   lack of objectivity, 116–117  
   personal experience used by, 117  
   strategies and tactics, 118–119  
 Qualitative data, 250  
 Qualitative methods, 252–253  
 Qualitative observations, 22–23  
 Qualitative research, 251–252, 252t  
 Qualitative studies, 29, 95  
 Quality adjusted life years, 341  
 Quality adjusted prosthetic years, 341–342, 345  
 Quality of results, 175–176  
 Quantitative observations, 22–23  
 Quantitative research, 250, 252t  
 Quasi-experimental designs, 236–237, 298–299  
 Quincunx, 168

## R

Random error, 164, 167, 279  
 Random experiment, 129  
 Random sampling, 250  
 Random variable, 130  
 Randomization, 275–276, 291–292  
 Randomized block test, 159t  
 Randomized controlled trials, 235–236, 291–292, 299–300  
 Range, 185  
 Ratio, 161, 170, 181  
 Ratio scale, 159t, 159–160  
 Rational appeal, 42  
 Rationality, 74  
 Rationalization trap, 48  
 Readers, 8  
 Real world research, 250–251  
 Realism, 21



Reasoning  
 by analogy, 87–88  
 types of, 8

Reciprocation, 51

Recognition, 15

Reconstruction experiment, 290

Red herring, 50

Reductionism, 122

Referee(s)  
 authors versus, 7–8  
 description of, 5–6

Refereed journals, 4, 60, 61b

Reference distribution, 133

Reflexology, 121

Regression  
 correlation versus, 261–264  
 description of, 118, 241

Regression models, 272

Relative frequency histogram, 131, 131f, 133, 134f, 136f, 141f

Relevance, 176–178

Reliability  
 definition of, 206  
 of measurements, 172, 206–214  
 validity versus, 207, 208f, 213

Reliability coefficient, 212

Repeatability index, 210

Repeated conjunction, 99

Repetition, 25, 28

Replication, 213

Reporting bias, 118

Reporting practices, 27

Representativeness heuristic, 322

Reproducibility  
 description of, 10, 18  
 lack of, 26  
 lack of incentives for, 26–27  
 openness and, 25–29  
 personal accountability in, 28  
 plans to enhance, 28

Reproducibility crisis  
 comments regarding, 28–29  
 defective reporting of  
 methodology, 27  
 hypercompetition, 27–28  
 lack of incentives, 26–27  
 multifactorial reasons for, 26–28  
 seriousness of, 26

Reproducibility ratio, 212

Research. *See also* Experiment(s).  
 discovery-based strategies, 235  
 experimental, 280–281  
 external validity of, 242–243  
 funding of, 18, 232–234  
 internal validity, 239–241  
 niche for, 233–235  
 non-experimental design of, 235  
 open access, 63  
 pre-experimental design of, 236  
 problems associated with, 242–243  
 qualitative, 251–252, 252t  
 quantitative, 250, 252t  
 quasi-experimental design of,  
 236–237  
 real world, 250–251

study designs, 238–239  
 subjects for, 238–239, 241, 285  
 variables, 286

Research proposal, 233, 235

Research question, 132

Residual sum of squares, 169

Residuals, 262–263

Resolution, 153–154

Response bias, 38, 307

Response variables, 263

Results. *See also* Data.  
 bias effects on, 166  
 comparisons, 176  
 confidence intervals used to  
 predict, 149  
 frequency of, 136  
 goodness of fit, 138–142  
 in illustrations. *See* Illustrations.  
 integrity, 173–175  
 location of, 136  
 presentation of, 173–196  
 quality of, 175–176  
 relevance of, 176–178  
 in tables, 186

Results section, of scientific papers,  
 37–38

Retaliation, 16

Retrospective cohort study, 238

Rhetoric  
 canons of, 41–46  
 decision making and, 40  
 dispositio, 44  
 elocutio, 44–45  
 inventio, 41–44  
 medieval versus contemporary,  
 43b  
 overview of, 40–41  
 in scientific papers, 23

Risk factors, 270

Risk of error, 133

**S**

Sackett's diagnostic tests for  
 causation, 105–106

Sample  
 definition of, 133  
 from normal distribution, 148–149  
 from population of known  
 dispersion, 147–148  
 from population whose dispersion  
 is not known "a priori," 149

Sample size, 148f, 265, 303–304

Sampling, 243, 250

Sampling bias, 307

Sampling space, 129

Scales  
 deceptive uses of, 192–194  
 interval, 159, 159t  
 logarithmic, 193  
 nominal, 157  
 ordinal, 157–159, 158t  
 purpose of, 157  
 ratio, 159t, 159–160

Scarcity, 53

Science, 2, 10

Scientific argument, 87–88

Scientific discovery, 87–88

Scientific judgments, 315

Scientific literature  
 citation analysis used to evaluate.  
*See* Citation analysis.  
 growth rate of, 3  
 meta-analysis of, 149  
 review of, 4

Scientific method  
 common sense and, 13  
 hypothesis, 23  
 objectivity. *See* Objectivity.  
 observations, 21–23  
 statistics and, 301  
 storybook version of, 20–25

Scientific misconduct, 28

Scientific papers  
 abstract, 34–36  
 acceptance date for, 34  
 active reading and, 32–33  
 authors of, 33–34  
 collaboration on, 16  
 comparisons in, 176  
 components of, 33–38  
 date of acceptance for, 34  
 date of submission for, 34  
 definition of, 3  
 discussion/conclusions section  
 of, 38  
 evaluator of, 10  
 importance of, 10  
 introduction section of, 36  
 materials and methods section of,  
 36–37  
 multi-authored, 16  
 objectivity in, 7  
 originality of, 9  
 publication of, 23  
 purposes of, 44, 173, 235  
 questions to ask about, 9–11  
 reproducibility of, 10  
 results section of, 37–38  
 rhetoric in, 23  
 style of, 23  
 submission date for, 34  
 suitability of, 4  
 systematic approach to analyzing,  
 2  
 title of, 33  
 truthfulness of, 9–10  
 writing style for, 7

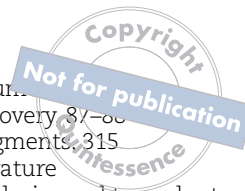
Scientific taste, 10

Scientific theories, 87

Scientific writing, 45, 316

Scientist behavior  
 aberrant, 17–18  
 collaboration, 16  
 everyday life, 14  
 motivators of, 15  
 objectivity. *See* Objectivity.  
 principles of, 14–15  
 traditional portrayal of, 15

Segmented regression analysis, 299



- Selection bias, 269, 306  
 Self-censorship, 52  
 Semi-formal logic, 77  
 Semi-log plot, 193–194  
 Sensitivity, 201b, 207t  
 Sensitivity analysis, 343–345  
 Sign test, 158t  
 Significance level, 133  
 Significant digits, 186  
 Silver amalgam replacement, 121  
 Simpson paradox, 181  
 Snow job, 51  
 SOAP format, 198  
 Soapbox effect, 50  
 Social validation, 52  
 Special pleading fallacy, 122  
 Specificity, 201b, 207b  
 Split-mouth design, 298  
 Square of opposition, 79, 79f  
 Standard(s)  
   changing of, 9  
   correlational experiments  
     involving humans, 269–270  
   of the field, 318  
 Standard deviation, 154, 170, 185  
 Standard error, 148, 185–186  
 Standard error of measurement, 212  
 Standard normal distribution, 147  
 Statistical adjustment, 264  
 Statistical data. *See* Data.  
 Statistical generalization, 302  
 Statistical inference  
   as inductive argument, 302–305  
   description of, 131  
   error, 137–138  
   features of, 138  
   hypothesis testing, 131–137  
   nonparametric, 140  
   one-tailed versus two-tailed tests,  
     137, 311  
 Statistical modeling, 270  
 Statistical power, 125  
 Statistical regression, 241  
 Statistical tests  
   differences in conditions, 323–324  
   nonparametric, 158t  
   types of, 158t  
   validity of, 155–157  
 Statistics  
   definition of, 128  
   descriptive, 128, 305  
   inferential, 128, 305  
   nonparametric, 140  
   ongoing development of, 150  
   reporting of results, 135  
   resources for, 143  
   rules for, 311  
   scientific method and, 301  
 Stereotype threat, 157  
 Stratified randomization, 291  
 Structured abstract, 35–36  
 “Structured discussion,” 38  
 Student distribution, 149  
 Subgroup analysis, 181  
 Subject heading searches, 66–67, 69  
 Subjective information, 198  
 Subjective probability, 129  
 Subjects, 238–239, 241, 285  
 Submission date, for scientific  
   papers, 34  
 Success fallacy, 124–125  
 Sufficient condition, 99–100  
 Sum of squares, 169  
 Suppressed premises, 82  
 Surrogate variables, 22  
 Surveys, 38, 238, 258–265  
 Syllogisms  
   categorical, 80–82  
   categorical statements, 79–80  
   definition of, 78–79  
   disjunctive, 82, 135  
   mixed hypothetical, 81  
   pure hypothetical, 81  
 Synthetic approach, 282  
 Systematic errors, 164, 165–166, 279  
 Systematic reviews, 3, 57–59, 58b
- T**  
 t distribution, 149  
 t test, 159t  
 Tables, 186  
 Target population, 305, 307  
 Technical quality, 175–176  
 Temporal proximity, 98  
 Temporality, 98–99  
 Temporary abandonment principle,  
   281  
 Testability of hypotheses, 93, 116  
 Tests. *See* Diagnostic tests and  
   measurements; Statistical  
   tests.  
 Textbook errors, 25  
 Theory, 87, 113  
 Threshold approach, to decision  
   analysis, 202, 202f  
 Title, of scientific papers, 33  
 Transfer method of research, 280  
 Transferability, 250–251  
 Treatment(s)  
   anecdotal evidence, 110  
   approaches to, 110–111  
   cause-effect reasoning and, 105  
   complex, 118  
   definition of, 200  
   ineffective treatments appearing  
     successful, 117–118  
   numerical method, 110  
   pathophysiological approach to,  
     110–111  
   personal experience about, 117  
   questions about, 111–112  
   variability of conditions, 118  
 Treatment allocation bias, 307  
 Triangulation, 251  
 Trimming, 17  
 True biologic variation, 171  
 True control, 278  
 True experimental design, 237,  
   298–299
- True-negative rate, 122  
 True-positive rate, 122  
 Truncation, 68–69  
 Trust, 16, 91  
 Tu quoque argument, 122  
 Tufte’s evaluative ratios for graphs,  
   189–190  
 Two-factor ANOVA, 159t  
 Two-tailed tests, 137, 311  
 Two-way ANOVA, 158t  
 Type I error, 138, 138t  
 Type II error, 137, 138t
- U**  
 UFO fallacy, 124, 307  
 Ullrich’s Index of Periodicals, 4  
 Uncertainty  
   description of, 149, 279–280,  
     337–339  
   judgments under, 319–323  
 Unidimensional data, 195–196  
 Uniform biologic response, 99  
 Uniform distribution, 145  
 Unintentional bias, 18–19  
 Units, 160–161, 309–310  
 Universalism, 14  
 Utility, health-state, 340–341  
 Utility value, expected, 341–344, 349
- V**  
 Valid, 85  
 Valid categorical syllogism, 80  
 Validity  
   external, 242–244, 298  
   internal. *See* Internal validity.  
   of experiments, 280  
   of statistical tests, 155–157  
   reliability and, 207, 208f, 213  
 Value judgments, 339  
 Variability  
   biologic, 154, 208  
   from examination equipment and  
     environment, 208  
   examiner, 209–214  
   reporting, 208–209  
 Variables  
   confounding, 101, 260, 263  
   continuous, 145  
   dependent, 274–275  
   discrete, 130, 145  
   independent, 274–275  
   outcome, 263  
   response, 263  
 Variances, 154, 171  
 Venn diagram, 129, 130f  
 Visual analog scale, 340f, 340–341  
 Volunteer bias, 306
- W**  
 Walton, Douglas, 77, 87  
 Washout period, 298

Wilcoxon rank sum test. See Mann-Whitney U test.  
Willingness to pay analysis, 345–347  
Wisdom of crowds, 24  
Withdrawal bias, 306  
Writing style, 7

WTP. See Willingness to pay analysis.

**Y**  
Yoked control, 293

**Z**  
z score, 146, 159t, 180

