Reappointment Packet
Russell Marcus
Hamilton College
January 2012

Research Materials

This file contains all of my publications since coming to Hamilton College; a copy of the proposal and acceptance letter for *Autonomy Platonism and the Indispensability Argument*, which is under contract at Lexington Books; the proposal for my co-edited philosophy of mathematics reader, as well as two works in progress. The table of contents here does not contain hyperlinks to locations within the document, but the bookmarks function for the pdf should assist your navigation.

Table of Contents

1. Proposal and acceptance letter for *Autonomy Platonism and the Indispensability Argument*
2. "Intrinsic Explanation and Field's Dispensabilist Strategy," forthcoming in *International Journal of Philosophical Studies*. A short version of the paper was presented on the main program at the APA Eastern Division Meeting, December 29, 2007; and at the Sydney-Tilburg Conference on Reduction and the Special Sciences, Tilburg University, The Netherlands, April 10, 2008 (both refereed).
3. "Structuralism, Indispensability, and the Access Problem," *Facta Philosophica* 9, 2007: 203-211.
4. "Indispensability Arguments in the Philosophy of Mathematics, The *Internet Encyclopedia of Philosophy*, First Posted: October 18, 2010.
5. Entries on Alonzo Church, David Hilbert, Leopold Löwenheim, Model Theory, Semantic Trees, Thoraf Skolem, Syntax, and Variable; in *Key Terms in Logic*, Jon Williamson and Federica Russo, Continuum Publishers, November 2010.
6. "The Explanatory Indispensability Argument," work in progress, presented on the main program at the APA Eastern Division Meeting, December 30, 2010; the 3rd Cambridge Graduate Conference on the Philosophy of Logic and Mathematics, January 17, 2010; and as a plenary talk at the First Colombian Conference in Logic, Epistemology, and Philosophy of Science, November 6, 2009 (all refereed).
7. Proposal for *An Historical Introduction to the Philosophy of Mathematics*
8. Review of Kenneth Harrelson, *The Ontological Argument from Descartes to Hegel*, *The APA Newsletter on Teaching Philosophy* 10.1, Fall 2010: 11-13.
9. "A Cooperative-Learning Lesson Using the *Objections and Replies*," *The APA Newsletter on Teaching Philosophy* 9.2, Spring 2010: 5-9.
10. "Embracing the Cartesian Circle," work in progress
11. "Observations on Cooperative-Learning Group Assignments," *The APA Newsletter on Teaching Philosophy* 9.2, Spring 2010: 2-5.

Autonomy Platonism and the Indispensability Argument

Russell Marcus
Department of Philosophy, Hamilton College
198 College Hill Road
Clinton NY 13323
rmarcus1@hamilton.edu
(315) 859-4056 (office)
(315) 381-3125 (home)

Autonomy Platonism and the Indispensability is a proposed research monograph. It will be appropriate for graduate students and other researchers in the philosophy of mathematics, as well as advanced undergraduates. It could be used in any course on contemporary philosophy of mathematics, and advanced courses in metaphysics.

The book will fill a void in the literature that has existed at least since the death of Jerrold Katz, who defended a version of autonomy platonism in his *Realistic Rationalism*. (I worked with Jerry on the early stages of my dissertation.) While the indispensability argument has nearly dominated research in the philosophy of mathematics over the last thirty years, and current interest in the argument is high, no one is writing about the argument from an autonomy platonist's perspective. My dissertation, *Numbers Without Science*, CUNY 2007, focused on the topic, and I have been working on this material since then.

Overview of the Argument

The central thesis of *Autonomy Platonism and the Indispensability Argument* is that autonomy platonism can break a stalemate in the debate over the indispensability argument by capturing the most desirable aspects of the two most important competing views.

Autonomy platonism is both a metaphysical and epistemological view about mathematics. On the metaphysical side, it is the claim that many mathematical theorems are true and that mathematical objects exist. On the epistemological side, it is the claim that our knowledge of mathematics does not derive from or depend on our knowledge of empirical science.

The indispensability argument in the philosophy of mathematics is also an argument for platonism about mathematics, an indispensability platonism. Following the work of Quine and Putnam, many platonists including Michael Resnik, Mark Colyvan, and Alan Baker, have defended the indispensability argument. These indispensabilists claim that we can not make sense of science without assuming the existence of mathematical objects.

There are two notable ways to oppose the platonist conclusion of the indispensability argument. The first way is to question its soundness, to deny that one must quantify over mathematical objects in one's best theory. Much recent work in the philosophy of mathematics, dating back to Hartry Field's 1980 *Science without Numbers*, has accepted the validity of the argument and focused on whether one can eliminate quantification over mathematical objects in physical theory or whether mathematical axioms must be included among those of our most sincere theories of the world.

Alternatively, one could deny the validity of the indispensability argument. Recent detractors of the argument claim that the inference from our uses of mathematics in science to their existence is fallacious. Almost all of the opponents of the indispensability argument, from Hartry Field to Joseph Melia and Mary Leng, have been anti-platonists, whether they deny its soundness or its validity.

The debates over the indispensability argument appear to be at a stalemate. On the indispensabilist's side, it is clear that no neat, first-order theory which eschews all mathematical axioms will suffice for all of current and future science. In contrast, the dispensabilist has reasonable hope of finding moderately attractive reformulations of large swaths of scientific theory.

Autonomy Platonism and the Indispensability Argument opposes the indispensability argument from the platonist's point of view. Autonomy platonism embraces both the indispensabilist's claim that we should take the references to mathematics in physical science seriously and the anti-indispensabilist's claim that the inference from mathematical references within science to the truth of mathematical claims is fallacious. Given that autonomy platonism can avoid the stalemate in the debate over the indispensability argument, I argue, we should posit the existence of mathematical facts, independent of the physical world, about which mathematical claims may be true or false.

The book is influenced by Jerrold Katz's work in linguistics; autonomy platonism is the natural companion to his intensionalist views. The indispensability argument is a reluctant platonist's argument, or even a nominalist's argument, and not the best argument for platonism. Moreover, the indispensability argument is at a stalemate and this book shows a way to break that stalemate.

Brief Outline of the Book

Part One. *Platonism: An Overview*

I present a general characterization of autonomy platonism, and trace its historical precedents. I show how the indispensability argument provides a promising response to the underlying problem of rectifying access to mathematical objects with standard theories of semantics and truth. I also show that autonomy platonism is consistent with a simple account of the applicability of mathematics, adapting and extending Mark Balaguer's account of applicability for plenitudinous platonism.

Part Two. *The Quinean Indispensability Argument*

Quine's indispensability argument, as I interpret it, is the strongest version of the argument. I show both how the argument is really a nominalist's argument and how Quine's version deflects recent so-called weaseling attacks on the argument from Joseph Melia and Mary Leng. The weasel argues that the role of mathematics in scientific theories is merely representational, and so our uses of mathematics are not to be taken literally.

Part Three. *The Putnamian Indispensability Argument*

There are weaker versions of the indispensability argument. Older versions are from Michael Resnik and Hilary Putnam. These versions are liable to weaseling responses.

Part Four. *The Explanatory Indispensability Argument*

The most recent version of the indispensability argument is due to Alan Baker and Mark Colyvan. Like the versions discussed in Part Three, the explanatory argument succumbs to weaseling attacks.

Part Five. *A Non-Indispensabilist Platonism*

We can draw two morals from the discussions of Parts Two through Four. First, the inference from our uses of mathematics in science to the truth of mathematical claims is, at the very least, not obviously valid. Second, scientific theories and explanations are more satisfying when they include references to mathematical objects. The autonomy platonist, by distinguishing evidence for mathematics from the presumption of truth for mathematical claims made within scientific theories, can embrace both of these morals.

Current Status of the Project, Summer 2011

I have completed most of the research for the book, though new material on the indispensability argument continues to appear rather frequently (as long as Mark Colyvan continues to breathe!). Parts One to Three and Part Five are based on some work from my 2007 dissertation, *Numbers without Science*.' More specifically, much of Parts One and Five come out of the conclusion of my dissertation; I have no firm plans to publish the rest of this material elsewhere, though I plan a longer, extended defense of autonomy platonism in the future.

Some material from Part One comes from my "Structuralism, Indispensability, and the Access Problem," *Facta Philosophica* 9, 2007: 203-211. Some of the exegetical material, especially in Parts Two and Three will rely on my *Internet Encyclopedia of Philosophy* article, "The Indispensability Argument in the Philosophy of Mathematics." But, the book would be a cleaner, streamlined version of the dissertation work, incorporating new insights and more recent research.

Part Four is new, and the basis for a paper I presented on the main program of the Eastern APA December 2010. Indeed, the idea for this book arose from my realization that I needed a longer format to properly contextualize my argument in that paper. An extended version of the APA paper is currently under consideration at a journal. Some of that material structures Parts Two and Three. An extended version of my work on Putnam's indispensability argument, Part Three, is in preparation for a separate journal submission.

Similar Books and Articles

Autonomy Platonism and the Indispensability Argument is intended as a contribution to the literature which includes the following books:

Azzouni, Jody. 2004. *Deflating Existential Consequence: A Case for Nominalism*. New York: Oxford University Press.

Balaguer, Mark. 1998. *Platonism and Anti-Platonism in Mathematics*. New York: Oxford University Press.

Colyvan, Mark. 2001. *The Indispensability of Mathematics*. Oxford University Press.

Leng, Mary. 2010. *Mathematics and Reality*. Oxford University Press.

Relevant recent articles in the area include:

Baker, Alan. 2005. "Are there Genuine Mathematical Explanations of Physical Phenomena?" *Mind:* 114: 223-238.

Bangu, Sorin Ioan. 2008. "Inference to the Best Explanation and Mathematical Realism." *Synthese* 160: 13-20.

Colyvan, Mark. 2010. "There's No Easy Road to Nominalism." *Mind* 119 (474): 285-306.

Mancosu, Paolo. 2008. "Explanation in Mathematics." *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/mathematics-explanation/>.

Melia, Joseph. 2000. "Weaseling Away the Indispensability Argument." *Mind* 109: 455-479.

Yablo, Stephen. 2005. "The Myth of the Seven." In *Fictionalism in Metaphysics*, M. Kalderon, ed., New York: Oxford University Press.

Timeline for Completion

My teaching responsibilities at Hamilton College and other projects will not leave me enough time to complete this project until my sabbatical year, 2012-13. I plan to take that full year, and finish the monograph by Summer 2013.

Possible Referees

Alan Baker: abaker1@swarthmore.edu
Associate Professor
Department of Philosophy
Papazian Hall 213
Swarthmore College
500 College Avenue
Swarthmore, PA 19081-1397

Mark Balaguer: mbalagu@calstatela.edu
Professor
Department of Philosophy
Cal State LA
5151 State University Drive
Los Angeles CA 90032

Mark Colyvan: mcolyvan@mail.usyd.edu.au
Professor of Philosophy
Director, Sydney Centre for the Foundations of Science
Department of Philosophy
A14 Main Quadrangle
University of Sydney
Sydney, NSW, 2006
Australia

Mary Leng: mcleng@liv.ac.uk
Department of Philosophy
University of Liverpool
7 Abercromby Square
Liverpool
L69 7WY U.K.

Paolo Mancosu: mancosu@socrates.berkeley.edu
Professor of Philosophy
Chair, Department of Philosophy
314 Moses Hall #2390
University of California
Berkeley, CA 94720-2390

Autonomy Platonism and the Indispensability Argument is not currently under consideration with any other publisher.

Lexington Books
A Wholly Owned Subsidiary of the Rowman & Littlefield Publishing Group, Inc.
4501 Forbes Boulevard, Suite 200 • Lanham, Maryland 20706

Offered 8/9/2011

PUBLISHING AGREEMENT

AUTHOR PREPARES CAMERA-READY MANUSCRIPT

Author: Russell Marcus

The Work: *Autonomy Platonism and the Indispensability Argument*

The above-named and undersigned (the Author) hereby agrees with Lexington Books (the Publisher) to the following terms and conditions in connection with its publication of the manuscript listed above (the Work):

1. AUTHOR SHALL GRANT TO THE PUBLISHER during the term of the United States copyright, and during any renewal or extension thereof, the exclusive right to publish and sell, including the right to permit others to publish and sell, the Work in any and all editions whether print, electronic, or audio, in all languages throughout the world.

2. AUTHOR SHALL DELIVER TO THE PUBLISHER at Author's expense no later than **August 1, 2013** (Delivery Date), two copies of the Work, including any photographs, tables, illustrations, or charts and bibliography or bibliographic essay, in camera-ready form, as well as all necessary permissions, licenses, releases, and consents. All graphics must be in camera-ready form. At least one month prior to the date above, Author shall send sample camera-ready pages for Publisher to review for format and quality before the entire text is delivered. Publisher shall have entire text proofread at no expense to Author. Author agrees to review and approve the proofread pages, to make requested corrections, and to supply any required indexes. When Author has made all necessary corrections, Author agrees to submit to Publisher (1) final camera-ready pages that have been printed on a laser printer with a resolution of not less than 600 dpi (dots per inch) or typeset professionally and (2) copies of the corrected electronic word processing files on disk for archiving purposes.

In order to be considered satisfactory, the Work must be factually accurate, original, and must acknowledge all intellectual debts. The Publisher may choose to engage scholars, consultants, or other experts in the topic covered by the Work to help determine its accuracy and originality. Any permissions necessary in order to reprint already published and/or copyrighted material quoted in the Work shall be obtained and paid for by Author.

If Author does not so deliver the Work to the Publisher in form and substance satisfactory to the Publisher in its judgment (which shall be final) by the Delivery Date specified or at another date agreed to by Publisher in writing (such time to be deemed of the essence), Publisher may terminate this agreement by notice to Author; such termination, if not caused by circumstances beyond control, shall be without prejudice to any other remedies Publisher may have for breach of contract.

3. THE PUBLISHER AGREES TO PUBLISH THE WORK at its expense in a manner and style and at a price that it determines. All decisions and details as to the editing and publication of the Work, including style,

edition or format, illustrations, time and manner of production, advertisement, and the number and free copies distributed will be left to the Publisher's sole discretion.

4. THE PUBLISHER AND THE AUTHOR SHALL TAKE SUCH STEPS as necessary on their respective parts to secure and preserve a valid United States copyright for the Work to be held in the Publisher's name. If the Publisher deems it desirable it may obtain copyright in other countries. The Publisher shall publish the Work in the United States in a manner that complies with the provisions of the Universal Copyright Convention.

5. IF THE COPYRIGHT OF THE WORK is infringed during the term of this agreement, then, upon notice to the effect by either the Author or the Publisher, both shall confer with regard to the infringement; if no mutually satisfactory arrangement is arrived at for joint action within ten days thereafter, either one shall have the right to bring an action or proceeding to enjoin the infringement, and for damages. If the proceeding is done jointly, the expense and recoveries, if any, shall be shared equally, and if there is no agreement to proceed jointly, either party shall have the right to go forward with an action or proceeding, bearing all the expenses and retaining any recovery. If the Publisher proceeds alone, the Author shall permit the action to be brought in his/her/their name and shall take all steps necessary, including the execution of such documents as may be required, to enable the Publisher to proceed.

6. (a) THE AUTHOR COVENANTS, WARRANTS, AND REPRESENTS that the Author is the sole owner of the Work and all the rights herein granted and has full right and power to make this Agreement; that the Work is not a violation of any copyright, proprietary or personal right; that the Author has not in any manner disposed of any of the rights herein granted to the Publisher or granted any rights adverse to or inconsistent therewith; nor are there any rights outstanding that would diminish, encumber, or impair the full enjoyment or exercise of the rights herein granted the Publisher; that no part of the Work is libelous, obscene or unlawful, or violates any right of privacy or any other right of a third party. In no event shall the Publisher be obligated to publish a work that, in its opinion, may subject it to any claims from any third party. Without limiting the foregoing the Author shall, if requested by the Publisher, make such changes in the Work as it shall deem necessary, but failure on the part of the Publisher to require or recommend any changes in the Work shall in no way affect the Author's obligation under this Paragraph.

6. (b) The Author agrees to hold harmless and indemnify the Publisher against any claim, demand, suit, action, proceeding, recovery, or expense of any nature whatsoever arising from or based upon any breach or alleged breach of the covenants, warranties or representations contained in subparagraph 6 (a) hereof or elsewhere in this agreement. In addition to all other remedies available to it, the Publisher may, in the event of any such claim, demand, suit, action, or proceeding, withhold advances and/or royalties due to the Author. The Publisher may at its sole discretion retain its own counsel, at its own expense, and the Publisher shall have sole discretion whether or not to settle or compromise any such claim, demand, suit, action, or proceeding brought against it. The warranties, representations and indemnities shall survive the termination of this Agreement.

7. (a) THE AUTHOR GRANTS TO THE PUBLISHER the exclusive right to license, sell or otherwise dispose of the following rights in the Work: publication or sale of the Work by publication of a reprint edition of the Work by another publisher in any edition or format whether print, electronic, or audio; editions prepared for book clubs; condensations; serializations in magazines or newspapers (whether in one or more installments and whether before or after book publication); translations; foreign English-language rights; electronic rights of every kind or nature; audio rights of every kind or nature; dramatic and motion picture rights and performance rights of every kind or nature, including but not by limited to, feature-length motion picture, television, cable, and internet rights as well as, film strips based on the story and film strips or motion pictures photographed directly from the book; phonograph and radio, audiotape, television, cable or other broadcasting rights; publication of the Work and selections therefrom in anthologies, compilations and digests; picturized book versions; microprint and microfilm versions; publication of the Work or derivative in part or combined with other works; publication in electronic multimedia versions, including the right to add sounds, images, or

graphics and to distribute such edition by any electronic means, method, device, process, or medium; and any and all other rights which may not be herein described.

In the case of each of the rights specified in this paragraph 7 (a) above, the net proceeds of the Work after third-party costs shall be shared equally between the Publisher and the Author and shall be paid on the regular royalty schedule specified in paragraph 9 (a) below.

7. (b) THE AUTHOR ALSO GRANTS TO THE PUBLISHER the exclusive right to license, sell or otherwise dispose of the following rights in the Work: electronic, including CD-ROM, CD-i, or online formats, mechanical, visual, and internet, intranet or web formats, or any other electronic means of reproduction, dissemination or distribution that exists or heretofore has not been established.

In the event the Publisher exercises the rights set forth in paragraphs 7 (a) or (b) rather than licensing them to third parties, the Publisher agrees to pay a royalty equal to the royalty specified below for any print edition.

8. (a) THE PUBLISHER SHALL PAY TO AUTHOR on each copy of print or electronic editions of the Work sold (except as provided in the following subparagraphs) a royalty of percentage of all sales or each edition or revision, domestic and foreign, as follows:

1 to 500 copies sold	0% of net sales
501 to 1,000 copies sold	5% of net sales
1,001 to 2,500 copies sold	7% of net sales
2,501 to 5,000 copies sold	8% of net sales
More than 5,000 copies sold	9% of net sales

The same schedule applies to hardcover, paperback, and electronic editions; however, hardcover, paperback, and electronic editions are computed separately.

8. (b) No royalty will be paid on copies furnished without charge or for review, advertising, sample, promotion or other similar purposes.

8. (c) Remainder sales will not be eligible for royalties, but prior to sale to a remainder dealer the Publisher will make reasonable efforts to notify the Author and accord the Author the opportunity to purchase all or part of such overstock at the remainder price.

8. (d) All sales subject to royalties under any of the provisions of this Paragraph 8 shall be computed net of returns.

8. (e) If more than one individual is designated as the Author, the royalty and all other monies due them shall be divided between or among them equally.

9. (a) FOLLOWING PUBLICATION PUBLISHER SHALL RENDER TO THE AUTHOR on June 30 of each year annual statements of account as of the preceding December 31. Each statement shall be accompanied by payment of all sums due thereon, including proceeds from the sale of any subsidiary rights. In reporting sales the Publisher may withhold a reasonable reserve for future returns. The Publisher may take credit for previous overpayments resulting from the return of books and for any other amounts due to the Publisher or chargeable against the royalties of the Author under the terms of this Agreement or any other agreement between the parties. If in any annual period the total payments due are less than $50.00, the Publisher may defer the rendering of statements and payments until such time as the sum of $50.00 or more shall be due. Payments will be made in U.S. dollars.

9. (b) The Author may, upon written request, examine the Publisher's books of account insofar as they relate to the Author's Work.

10. **THE PUBLISHER MAY PUBLISH** or permit others to publish or broadcast without charge and without royalty such selections from the Work for publicity purposes as may, in the Publisher's opinion, benefit the sale of the Work. The Publisher also may authorize publication of the Work without charge and without royalty in Braille or by any other method primarily designed for the physically handicapped.

11. **THE AUTHOR AGREES** that for the duration of this contract, both prior to the publication of the Work and as long as the Work remains in print, he/she/they will not author, co-author, or serve as general editor for a competing work or one that would substantially inhibit the sales of the Work named herein.

12. (a) **ON PUBLICATION OF THE WORK,** Publisher shall deliver to the Author six (6) copies without charge and the Author shall have the right to purchase further copies (but not for resale) at a discount of 50% from the Publisher's retail price, plus shipping and handling. These sales shall be royalty exclusive. Copies for resale may be purchased at a 25% discount, plus shipping and handling. The Author may purchase other titles on the Publisher's list at a discount of 25% from the retail price, plus shipping and handling.

12. (b) In the case of an edited volume, each editor of the Work shall receive two (2) copies without charge. Each contributor to the Work shall receive one copy of the Work without charge. (Editor must supply all current addresses in order for copies to be sent.) Editor(s) and contributors shall have the right to purchase further copies (but not for resale) at a discount of 50% from the retail price, plus shipping and handling. These sales shall be royalty exclusive. Copies for resale may be purchased at a 25% discount, plus shipping and handling.

13. **IN THE EVENT THE PUBLISHER FAILS TO KEEP THE WORK** in print (and for all purposes of this paragraph the Work shall be considered to be in print if it is on sale by Publisher in any edition or if it is on sale in any edition licensed by Publisher or available electronically or in a print-on-demand edition during the term of this Agreement) and Author makes a written request of Publisher to keep the Work in print, the Publisher shall, within sixty days after the receipt of the request, notify the Author in writing whether or not the Publisher intends to do so, and if Publisher elects to do so. If the Publisher elects to do so, the Publisher shall have six months thereafter in which to comply. If the Publisher fails to comply (unless the failure is due to circumstances beyond the Publisher's control), or if the Publisher does not desire to keep the Author's work in print, then this Agreement shall terminate and all rights shall revert to Author.

14. **AUTHOR SHALL EXECUTE AND DELIVER** whatever documents and assignments of copyright or other papers as may be necessary in our opinion to fulfill the terms and intent of this agreement.

15. **IN ORDER TO KEEP THE WORK UP TO DATE** the Author shall, if requested by the Publisher, revise the Work for new editions while this Agreement is in effect and shall supply any new matter that may be needed to that end. All terms of this Agreement shall apply to those new editions or revisions, including payment of royalty at the original schedule of rates provided for the first edition. In the event the Author is unwilling or unable by reason of death or otherwise to revise the Work or supply new matter in a form satisfactory to the Publisher, the Publisher may select some other person to edit or revise the Work or supply new matter and may deduct the expense thereof (including royalties to such other person) from the royalties payable to the Author and in such revised edition may describe such person as editor or co-author.

16. **EXCEPT FOR** loss or damage due to the Publisher's own negligence, the Publisher shall not be responsible for loss or damage of any of Author's property, including the manuscript of the Work, and liability for any such loss or damage shall in no event exceed the amount payable to the Publisher under any insurance carried by the Publisher covering such loss.

17. THIS AGREEMENT shall be deemed made in, and shall be in all respects interpreted, construed and governed by the laws of the State of Maryland, and the parties agree to litigate any controversy, claim or dispute arising out of or in connection with this Agreement or the breach thereof solely in the Maryland courts, which the parties agree shall have exclusive jurisdiction to which the parties hereby submit.

18. ANY NOTICE CALLED FOR in this Agreement shall be sent by mail to the Author or to the Publisher at the address set forth above. Either party may designate a different address by notice given in the same manner.

19. THE WAIVER of a breach of, or of a default under, any of the terms of this Agreement shall not be construed a waiver of any subsequent breach or default. No waiver or modification of this Agreement shall be valid unless in writing and signed by the party to be charged.

20. THIS AGREEMENT shall inure to the benefit of and be binding upon the Author, his/her heirs, personal representatives, and assigns and upon the Publisher, its successors, and assigns. The Author shall not have the right to assign this Agreement without the Publisher's written approval except that the Author shall have the right to assign any monies due the Author without such approval upon prior written notice to the Publisher. The Publisher shall have the right, without approval of the Author, to assign this Agreement.

21. ALL RIGHTS AND LICENSES granted or assigned by the Publisher, pursuant to this Agreement, to any division, affiliate, subsidiary, or parent company or successor of Lexington Books, shall be construed as though the Work was still in the Publisher's hands and all parts of this Agreement would still be in full force and effect.

22. IF MORE THAN ONE INDIVIDUAL IS DESIGNATED AS AUTHOR of the Work, then all obligations and indemnifications are joint and several.

23. THIS AGREEMENT CONSTITUTES THE WHOLE AGREEMENT between the Author and the Publisher and supersedes all preexisting agreements or arrangements, written or oral, between the parties relating to the subject matter of this Agreement.

24. THIS CONTRACT OFFER is valid if signed and returned within one hundred and eighty days of the offered date on page one.

by Julie E Kirsch 8/29/11

Julie E. Kirsch, Vice President/Publisher
Lexington Books

Date

Accepted and Agreed to: ______________________ 8-22-11

Russell Marcus Date

Date of Birth: 4-13-66 Citizenship: USA

Social Security Number: 130-46-4117

Preferred mailing address:

23 FOUNTAIN STREET

CLINTON NY 13323

Phone number: 315-381-3125

Email address: RMARCUS1@HAMILTON.EDU

Intrinsic Explanation and Field's Dispensabilist Strategy

Forthcoming in *International Journal of Philosophical Studies*

Russell Marcus, Ph.D.

Department of Philosophy, Hamilton College

198 College Hill Road

Clinton NY 13323

rmarcus1@hamilton.edu

(315) 859-4056 (office)

(315) 381-3125 (home)

January 3, 2012

Abstract:

Hartry Field defended the importance of his nominalist reformulation of Newtonian Gravitational Theory, as a response to the indispensability argument, on the basis of a general principle of intrinsic explanation. In this paper, I argue that this principle is not sufficiently defensible, and can not do the work for which Field uses it. I argue first that the model for Field's reformulation, Hilbert's axiomatization of Euclidean geometry, can be understood without appealing to the principle. Second, I argue that our desires to unify our theories and explanations undermines Field's principle. Third, the claim that extrinsic theories seem like magic is, in this case, really just a demand for an account of the applications of mathematics in science. Finally, even if we were to accept the principle, it would not favor the fictionalism that motivates Field's argument, since the indispensabilist's mathematical objects are actually intrinsic to scientific theory.

Keywords:

Philosophy of Mathematics, Indispensability Argument, Hartry Field, Intrinsic Explanation

§1: Overview

Quine argued that we should believe that mathematical objects exist because of their indispensable uses in scientific theory. Hartry Field rejects Quine's argument, arguing that we can reformulate science without referring to mathematical objects. Field provided a precedental reformulation of Newtonian Gravitational Theory (NGT) which has been refined, improved, and extended in the years since his original monograph. In this paper, I argue that Field's impressive construction and its extension do not impugn Quine's argument in the way that Field alleges that they do. I do not defend the indispensability argument. I merely attempt to undermine Field's influential line of criticism.

Field's reformulation of NGT is simply a formal construction. Field argues for its relevance in a defense of nominalism on the basis of a principle of intrinsic explanation. I argue that this principle is not sufficiently motivated or defensible, and that it can not do the work for which Field uses it. I start with the relevant background, in §2, and a discussion of Field's principle, in §3. In §4-§6, I present and reject Field's arguments for that principle. In §7, I show that even accepting Field's principle would not lead to his nominalist, or fictionalist, conclusion.

§2: Quine's Indispensability Argument and the Dispensabilist Response

Quine never presented a detailed indispensability argument, though he alluded to one in many places. I interpret Quine's argument as follows.

QI
- QI1. We should believe the single, holistic theory which best accounts for our sense experience.
- QI2. If we believe a theory, we must believe in its ontological commitments.
- QI3. The ontological commitments of any theory are the objects over which that theory first-order quantifies.
- QI4. The theory which best accounts for our sense experience first-order quantifies over mathematical objects.
- QIC. We should believe that mathematical objects exist.[1]

[1] See Quines 1939, 1948, 1951, 1955, 1958, 1960, 1978, and 1986. For other versions of the indispensability argument, see Putnam 1971 (the success argument); Resnik 1997: §3.3 (the pragmatic indispensability argument); and Mancosu 2008: §3.2 (the explanatory indispensability argument). I focus on Quine's argument because Field's response is directed at it.

An instrumentalist who believes that our uses of mathematics in science do not commit us to the existence of mathematical objects, may deny either QI1 or QI2, or both.[2] Regarding QI1, there is some debate over whether we should believe our best theories. Regarding QI2, one might interpret some of a theory's references fictionally. I return briefly to instrumentalist responses to QI in §7 of this paper.

Quine's procedure for determining the ontological commitments of theories, QI3, is less controversial than QI1-2, and may be taken as definitional. Still, instrumentalists who dismiss QI should be prepared to defend alternative criteria for determining their ontological commitments. One alternative to QI3 would be to adopt an eleatic principle on which the ontological commitments of any theory are, approximately, those objects in the causal realm.[3]

Debate over QI has focused mostly on QI4. To oppose QI4, Field provided two synthetic reformulations of NGT, replacing the standard analytic version of the theory, which relies on real numbers and their relations, with theories based on physical geometry. A second-order reformulation replaced quantification over mathematical objects with quantification over space-time points. A first-order reformulation referred instead to space-time regions. There are technical questions about whether Field's reformulations are adequate for NGT. Field mostly ceded the second-order reformulation, due to problems involving incompleteness.[4] The first-order version, using Quine's canonical language, is a more appropriate response to QI anyway. There are also questions about whether analogous strategies are available for other current and future theories.[5] I put these questions

[2] See Carnap 1950; and Melia 2000, Azzouni 2004, and Leng 2005 for more recent defenses of instrumentalism in response to QI.

[3] Contemporary discussions of the eleatic principle trace mainly to David Armstrong's work. Armstrong sometimes focuses on causation (see Armstrong 1978b: 46), at other time son spatio-temporal location (see Armstrong 1978a: 126). Other formulations are found in Oddie 1982: 286; Azzouni 2004: 150; and Field 1989: 68.

[4] See Shapiro 1983, Field 1985b, and Field 1990.

[5] Burgess and Rosen 1997 elegantly collects the slew of reformulation strategies published in the wake of Field's monograph. See especially the construction at §IIA in the spirit of Field's original work. Most reformulations replace mathematical references with modal ones.

aside, for this paper, and suppose that reformulations in the spirit of Field's construction are available for our best theories.

My concern in this paper is whether such reformulations are better theories than standard ones, for the purposes of QI. The superiority of dispensabilist reformulations is important because the indispensability argument relies on the claim, at QI1-2, that we find our ontological commitments in our best theory. Field defends his reformulation on the basis of a principle of intrinsic explanation. I argue that this principle is false, and that the standard theory is preferable to its dispensabilist counterpart. Thus, I reject Field's claim that QI4 is false, not because reference to mathematical objects is ineliminable from science, but because the reformulated theories which eliminate quantification over mathematical objects are not our best theories.

The value of dispensabilist reformulations has been questioned before. Pincock 2007 argues that the standard theory is better confirmed. To construct representation theorems which demonstrate that a reformulation is adequate, the dispensabilist adopts axioms about the physical world and its properties. For example, to measure mass or temperature, Field assumes the existence of spatio-temporal regions or points, and orderings among them, to do the work that connected sets of real numbers do in the standard theory.[6] But, Pincock argues, those assumptions about the physical world are not as well confirmed as the corresponding mathematical axioms and mappings between physical and mathematical structures.

Against Pincock's claim, even if the dispensabilist's axioms are less well confirmed than the mathematical axioms they replace, they may derive some measure of confirmation from their adequacy. Furthermore, the dispensabilist reformulation, eschewing mathematical objects, makes fewer commitments. It is not clear how to balance the virtue of having fewer commitments with the benefit of having a greater degree of confirmation.

Burgess and Rosen 1997 argue that a better theory should be publishable in scientific journals, and adopted by working scientists; since dispensabilist reformulations are not preferred by

---

[6] See Field 1980, Chapter 7. See Field 1985a for his arguments for a substantivalist interpretation of space-time.

practicing scientists, they are no better. This is a wrong way to measure the value of a theory. The practicing scientist wants a useful theory to produce and replicate empirical results. The scientist is mainly unconcerned with ontological commitments. Field's defense of his reformulation correctly emphasizes concerns about ontological commitments. Field shows that it is reasonable, even preferable, to continue using our standard theories, even if we do not really believe that the mathematical objects to which they refer exist. Dismissing concerns about the ontological commitments of our theories, as Burgess and Rosen do, ignores the questions raised by QI about whether scientific theories must quantify over mathematical objects.[7]

Much of the debate over whether dispensabilist reformulations are better than their standard counterparts has focused on their attractiveness. Field uses attractiveness as a criterion for acceptable reformulations in his original work.[8] But, attractiveness is a vague and malleable criterion. One might find a theory attractive based on its strength, simplicity, or explanatory power. It is unclear how to balance such considerations. "Of course, it is a deep and difficult question how the various attributes that contribute towards a theory's attractiveness ought to be spelled out, and how these attributes are to be independently measured and weighed against each other" (Melia 2000: 472).

Mark Colyvan, arguing that the standard theory is more attractive than Field's reformulation, mentions the unification achieved by the standard theory, and its boldness, simplicity, and predictive powers.[9] I believe that Colyvan's argument can be made more precise. This paper pursues and extends Colyvan's argument, criticizing Field's own criterion for attractiveness, his principle of intrinsic explanation. I present a specific explanation of why Field's reformulation is not a better or more attractive theory than the standard one.

§3: Intrinsic and Extrinsic Explanations and Theories

---

[7] On Burgess and Rosen's suggestion: "While entertaining as rhetorical flourishes, such demands leave a serious explanatory gap..." (Pincock 2007: 255).

[8] See Field 1980: viii, 8 and 41.

[9] See Colyvan 1999 and Colyvan 2001: §4.3.

Field defends his reformulation by appealing to a general preference for intrinsic explanations over extrinsic ones.

> If in explaining the behavior of a physical system, one formulates one's explanation in terms of relations between physical things and numbers, then the explanation is what I would call an *extrinsic* one. It is extrinsic because the role of the numbers is simply to serve as labels for some of the features of the physical system: there is no pretence that the properties of the numbers influence the physical system whose behavior is being explained. (The explanation would be equally extrinsic if it referred to *non-mathematical* entities that served merely as labels...) (Field 1985a: 192-3).

Field uses 'intrinsic' and 'extrinsic' to apply to entities, theories, and explanations. The application to entities is basic, and he classifies explanations and theories depending on the types of objects used. Explanations and theories are intrinsic if they make no demand for extrinsic objects.

According to Field, numbers are extrinsic to physics, while physical objects and space-time regions are intrinsic.[10] Numbers are extrinsic to geometry, too, while line segments and their properties are intrinsic. The application of 'intrinsic' within mathematics proper raises several questions about the relationships among mathematical theories. Are real numbers intrinsic or extrinsic to the theory of natural numbers? Are sets extrinsic to category theory? Are topological spaces extrinsic to Euclidean geometry?

Similar questions can be asked purely within empirical science. Are biological or psychological predicates extrinsic to physics? The objects of physics could be considered extrinsic to the special sciences, especially if there are emergent properties in those sciences. In fact, the classification of objects or theories as intrinsic or extrinsic seems suspiciously flexible. Consider how an Aristotelian would deem terrestrial objects as objects extrinsic to theories about planets and stars.

These questions about Field's principle within either mathematics or empirical science should make us wary of the commonsense intrinsic/extrinsic distinction. But, since my goal at this point is just to illustrate Field's distinction, and since my concern in this paper is with the relationship of mathematics to physical theories, I shall put them aside.

---

[10] Joseph Melia argues that space-time points are actually extrinsic to physical theories; see Melia 1998: 65-7.

Field presumes that mathematical objects do not influence physical systems to classify them as extrinsic to physical theory. "If, as at first blush appears to be the case, we need to invoke some real numbers...in our explanation of why the moon follows the path that it does, it isn't because we think that the real number plays a role as a *cause* of the moon's moving that way..." (Field 1980: 43).

Field's preference for intrinsic explanations is a broad methodological principle.

> Extrinsic explanations are often quite useful. But it seems to me that whenever one has an extrinsic explanation, one wants an intrinsic explanation that underlies it; one wants to be able to explain the behavior of the physical system *in terms of the intrinsic features of that system*, without invoking extrinsic entities (whether mathematical or non-mathematical) whose properties are irrelevant to the behavior of the system being explained). If one cannot do this, then it seems rather like magic that the extrinsic explanation works. (Field 1985a: 193; see also Field 1980: 44 and Field 1989: 18-9).

Call this principle PIE: we should prefer intrinsic explanations over extrinsic ones, when they are available. PIE is supposed to account for Field's preference for synthetic physical theories over analytic ones. PIE also supports Field's argument for substantivalist space-time (since relationalist theories require references to extrinsic real numbers) and explains his hostility to modal reformulations of science (since modal properties are extrinsic to physics).

Field's focus on intrinsic explanations, rather than intrinsic theories, might seem a bit puzzling. His project is clearly a response to Quine's indispensability argument, which is formulated in terms of theories because of Quine's demand that we find our ontological commitments in our best theory. To reformulate the indispensability argument in terms of explanations would force the indispensabilist to argue that we determine our commitments by consulting our best explanations. Though recent work by Colyvan and Alan Baker develops an explanatory indispensability argument, that argument is not Quine's argument, nor is it the argument to which Field is responding. I shall not pursue it here.[11]

If one thinks that scientific explanations are exhausted by the applications of our most austere scientific theories to sets of initial conditions, then there is no significant difference between appeals

[11] See Mancosu 2008, §3.2, for a formulation of the explanatory argument, and Baker 2005, Colyvan 2001, Colyvan forthcoming, and Lyon and Colyvan 2007 for defenses of the argument.

to explanations and theories. For traditional covering-law analyses of explanation, we need not worry whether Field's principle is made in terms of explanations or theories. Indeed, Field seems to have such a model of explanation in mind.

> What we must do is make a bet on how best to achieve a satisfactory overall view of the place of mathematics in the world... My tentative bet is that we would do better to try to show that the explanatory role of mathematical entities is not what is superficially appears to be; and the most convincing way to do that would be to show that there are some fairly general strategies that can be employed to purge theories of all reference to mathematical entities (Field 1989: 18; see also fn 15).

Furthermore, Field says that an explanation is, "A relatively simple non-*ad hoc* body of principles from which [the phenomena] follow" (Field 1989: 15).

In contrast, one might believe that criteria for good explanations are different from criteria for good theories, especially when theories are used for revealing ontological commitments. One might, say, wish that explanations be perspicuous. If so, one could not prefer Field's reformulation of NGT to the standard theory. Indeed Field's reformulation is imperspicuous, and hardly recognizable as NGT. It would be impossible to use, which is why he attempts to establish that mathematical theories are conservative over nominalist physical ones.[12] If we were to adjust QI to focus on explanation in this sense, a preference for intrinsic explanations could not support a dispensabilist reformulation. I will focus on PIE in the sense that I believe Field intended. In this standard sense, explanatory power is an important theoretical virtue.

§4: Field's Motivation for PIE: Hilbert's Intrinsic Geometry

I discern three arguments for PIE in Field's work. There is the 'magic' argument mentioned in the previous section, which I assess in §6. Field also relies on an implicit Okhamist argument, which I consider in §5. Lastly, Field argues that the importance of Hilbert's 1899 reformulation of

---

[12] If a mathematical theory is conservative over a nominalist physical theory, then we can use the mathematics to facilitate derivations in the physical theory with assurance that we will not derive any unacceptable empirical consequences. The conservativeness of mathematics would assure us that Field's reformulation need have no consequences for working scientists.

Euclidean geometry, which inspired Field's project, can be explained by PIE: Hilbert's axiomatization is superior because it is an intrinsic theory.[13] In this section, I argue that we can understand the success of Hilbert's axiomatization without adopting PIE as a general principle which supports Field's reformulation as a best theory for the purposes of QI.

The projects of axiomatizing mathematics in the late nineteenth century were motivated by diverse factors, two of which stand out: the oddities of transfinite set theory, and the development of non-Euclidean geometries.[14] In both cases, traditional mathematical ontology was contentiously extended without obvious inconsistency. Rigor in the form of axiomatic foundations was sought to put the controversial new theories on firm ground.

Since development of analysis in the seventeenth century, formulations of Euclidean geometry had used real numbers to represent lengths of line segments and triples of real numbers to represent points. Hilbert's new axiomatization referred to regions of geometric space in lieu of real numbers, and used the geometric properties of betweenness, segment congruence, and angle congruence in the way that real numbers, and their ordering, were used in analytic versions. Hilbert constructed representation and uniqueness theorems which assured the adequacy of his so-called synthetic theory.

We can understand why Hilbert would prefer a synthetic geometry over analytic versions without appealing to PIE. Here is what Hilbert says about his motivation:

> I wanted to make it possible to understand those geometrical propositions that I regard as the most important results of geometric inquiries: that the parallel axiom is not a consequence of the other axioms, and similarly Archimedes' axiom, etc. I wanted to answer the question whether it is possible to prove the proposition that in two identical rectangles with an identical base line the sides must also be identical, or whether as in Euclid this proposition is a new postulate. I wanted to make it possible to understand and answer such questions as why the sum of the angles in a triangle is equal to two right angles and how this fact is connected with the parallel axiom..." (Frege 1980: 38-9).

[13] On Hilbert's influence, see Field 1980, Chapter 3.

[14] Hilbert mentions both in a letter to Frege on his motivation for axiomatizing geometry (Frege 1980: Letter IV/4).

Hilbert's account of his motivations make it clear that he wanted to clarify relations within geometry. By relying on geometric relationships to explain geometric phenomena, he avoided worries about the consistency of analysis in addition to the worries about geometry. Hilbert thought that his axiomatization better explained geometric entailments.

We can best interpret Hilbert's motivation as purely mathematical, rather than ontological. He makes no suggestion that his new theory is better for the purposes of revealing ontology, which is how Field uses his formulation of NGT. Hilbert devised his axiomatization well before Hempel's work on scientific explanation, which linked explanation with formal theories, and well before Quine's work on ontological commitment, which linked formal theories with commitments. Furthermore, there are no benefits of parsimony arising from Hilbert's work. He only shows that real numbers are avoidable in the axiomatization of geometry. His project was not intended to eliminate commitments to numbers.

Regarding more general ontological questions, our main worry within mathematics is antinomy, not parsimony. Worries about antinomy within analysis motivated the arithmetization project of Cauchy, Weierstrass, Dedekind, and others, as well as Hilbert's axiomatization of geometry. But, contradictions may be more easily discovered in larger, more comprehensive theories than in smaller, more isolated ones. The superiority of Hilbert's axiomatization for the purposes of revealing geometric relations does not entail its superiority in constructing proofs and discovering contradictions. For a recently well-worn example, consider that Fermat's theorem, a number-theoretic claim, was proved by mapping formulas to topological spaces, after hundreds of years of more direct, intrinsic attempts to prove it.

Different axiomatizations serve different purposes. For ontological purposes, we are interested in the conjunction of Hilbert's construction with analysis, which maps geometric structures onto those of number theory. We need not invoke a general principle, PIE, to explain the utility of Hilbert's reformulation, and insist that his intrinsic, synthetic theory is superior to the analytic formulation. While Hilbert's axioms emphasize narrow geometric relations, they omit broader, edifying relations between analysis and geometry.

§5: Unification and Parsimony

Since we need not appeal to PIE to see the virtues of Hilbert's reformulation, Field's response to QI loses some of its motivation. More importantly, PIE, which seeks to isolate theories and explanations according to their intrinsic elements, conflicts with our general preference to unify theories, revealing connections among diverse disciplines. A comprehensive theory simplifies, by showing how different commitments cohere. As Michael Devitt argues, ontology is not to be found in isolated theories. "The best ontology will be that of the best unified science" (Devitt 1984: §4.9; see also §7.8).

For examples of the virtues of extrinsic theories within mathematics, consider how the fundamental theorem of calculus bridges geometry and algebra. Algebra could be seen as extrinsic to geometry, but uniting the two theories yields a more comprehensive, and more fruitful, theory. Also, we can prove more in an extrinsic second-order theory than we can in an incomplete first-order theory, like first-order arithmetic, itself.

The many precedents for unification in science include Newtonian Gravitational Theory unifying Kepler's celestial mechanics with Galileo's terrestrial mechanics, and Maxwell's electrodynamics unifying electrical and magnetic theories with optics. Consider how welcome bridge laws between physics and chemistry or biology would be. Indeed, unification may be central to our notions of scientific explanation. All of the most promising accounts of scientific explanation emphasize unification. Covering laws are preferred if they are broader. Causal accounts seek fewer, more unifying causes. On Kitcher's unificationist account, it is essential to an explanation that it unify a range of disparate phenomena. Kitcher even argues that unification is the underlying principle which the covering-law model was intended to capture.[15]

Of course, many good explanations appeal only to isolated portions of our theories. For example, even if we presume that mental-state predicates are somehow reducible to physical ones, a psychological explanation may not need to appeal to any physical principles. But, such limited

[15] See Kitcher 1981: 508.

explanations do not conflict with our broader presumption toward unification. Even a dualist would have to appreciate the development of bridge laws between cognitive and physical sciences. It would be implausible for Field to reject the general desire for unification, and there is no evidence that he does. But granted that we seek unifying explanations and theories, Field's general principle of intrinsic explanation seems difficult to defend.

One might wonder if the unification of mathematics and physics is a special case which resists a general preference for unification. Such resistance might be supported, say, by the observation that mathematical objects are causally independent of the physical world. Does our preference for unification prevail? Or, do the differences between mathematics and physical science entail that we should prefer scientific theories which eschew extrinsic mathematical objects? In the latter case, our preference for intrinsically isolating physics from mathematics is a specific case, not a corollary of a general principle. Even if one agreed with Field that his reformulation of NGT were a preferable, more attractive theory, that preference would not derive from a general principle of intrinsic explanation.

In the former case, we are left to wonder whether a limited principle of intrinsic explanation, in this particular case, supports a preference for Field's reformulation. Colyvan directly tackled the question of whether Field's reformulation is more attractive than standard NGT. He argued that unifying mathematics and science leads to a preferable theory. "Mathematics contributes to the unification and boldness of the physical theory in question, and therefore *is* supported by well-recognised principles of scientific theory choice" (Colyvan 2001: 81).

Colyvan provides three examples. First, the introduction of complex numbers as missing solutions to quadratic differential equations simplifies mathematics, since we need not wonder why some quadratic equations have only one, or even no, root. It unifies exponential and trigonometric functions, and any scientific theory which uses such functions. Second, Dirac predicted the existence of positrons by relying on the mathematical solutions to his eponymous equation in relativistic quantum mechanics; positrons were not experimentally verified for another five years. The unification of mathematics and physics allowed for faster scientific progress. Lastly, the Lorentz transformations were initially derived as an account of the failure of the Michelson-Morley experiment intended to provide evidence of the ether. Lorentz, who was in the grip of a false scientific theory, nevertheless developed equations which were later derived from a better scientific theory, viz. special relativity. Without the underlying extrinsic mathematics, it is difficult to see how Lorentz could have developed his equations.

Colyvan's examples illustrate how, in the absence of an over-riding principle, being intrinsic is just one among many characteristics to be weighed when evaluating the attractiveness of a theory. The principle of intrinsic explanation seems especially disfavored when applied specifically to the mathematics used in science, i.e. in the specific case on which QI depends. The unification of mathematics with physics yields a simpler and more powerful theory, a point which Field grants by arguing for conservativeness. And, the isolation of scientific theory from mathematics, especially on the basis of a dispensabilist reformulation, denies important relations among mathematical and physical objects. For a simple example, it is a mathematical property of a three-membered set that it has exactly three two-membered subsets. Applying this property, we can account for why we can, with a red marble, a blue marble and a green marble, form exactly three different-looking pairs of marbles.

The ability of a theory to unify disparate phenomena is only one factor among several that we use to evaluate theories. Others include strength, simplicity, fruitfulness, perspicuity, and parsimony. Complete lists are difficult to formulate. Field's argument is that we should include on such a list whether a theory is intrinsic, and I have argued that our desire to unify theories is more important. Still, the specific case at hand, the one for which Field invokes PIE, is whether standard (extrinsic) science is preferable to Field's nominalized (intrinsic) theory. The only obvious theoretical virtue of the nominalized theory is parsimony.[16]

Desire for parsimony proceeds from a principle applicable to the concrete objects posited by

[16] The nominalist theory may also be used in an account of the applicability of mathematics in empirical science, though the platonist can use any nominalist account just as well, so does not serve to distinguish the nominalist from the platonist.

our best theories: do not multiply physical entities without good reason. When constructing scientific theories, it is important not to posit more in the world than that which accounts for the phenomena.

It is an open question whether principles of parsimony should apply to the mathematical objects used in standard formulations of scientific theories. In mathematics proper, parsimony is not the most important theoretical virtue. In contrast to the natural scientist, the mathematician explores his universe with a desire to multiply entities. In mathematics, it is a virtue to be plenitudinous, as long as we avoid antimony. Once we have admitted abstracta into our ontology, we do not run out of room. Set theorists proudly present discoveries of distinct new cardinals. Kripke models for modal logic have ameliorated mathematical worries about modality without resolving persistent philosophical worries about possible worlds. The belief that principles of parsimony are applied differently in mathematics is also a basis for Mark Balaguer's plenitudinous platonism on which every consistent set of mathematical axioms truly describes a mathematical universe. Worries about the introduction of new mathematical entities, as with complex numbers, or transfinites, tend to focus mainly on their consistency, or the rigor with which they are introduced.

PIE would reduce the ontology of scientific theories at the expense of perspicuity, explanatory power, fruitfulness, and coherence with other theories. And it is not even clear that the reduced ontology is preferable. Burgess and Rosen argue against reduced mathematical ontology as a theoretic virtue. "It is at least very difficult to find any unequivocal historical or other evidence of the importance of economy of abstract ontology as a scientific standard for the evaluation of theories" (Burgess and Rosen 1997: 206).

The indispensability argument is alluring since it seems to provide a framework on which mathematical nominalists and platonists can agree.[17] Like QI, PIE was intended as a non-question-begging approach to the nominalist/platonist debate. But, the old debate remains.

---

[17] "On the one hand, the indispensability argument sides with nominalists in avoiding any presupposition that mathematical statements are intrinsically privileged. On the other hand, the argument sides with Platonists in taking mathematical statements at face value, as making genuine ontological claims... This evenhandedness is an important strength of the indispensability argument..." (Baker 2003: 50).

§6: The 'Magic' Argument

I have argued that two of Field's three arguments for PIE (its ability to explain the value of Hilbert's geometry and a general preference for parsimony over unification) are unsuccessful. Lastly, in defending his general principle, Field granted the utility of extrinsic explanations, but argued that they seem like magic if there is no underlying intrinsic explanation presupposed. Field's claim is that explanations of physical phenomena should be possible which refer only to entities which are active in producing those phenomena.

The obvious defense of the general demand for intrinsic theories comes from linking theories with ontological commitment, as Quine does. We want our theories to refer only to relevant objects in order to avoid errant commitments. But if we have the mathematical commitments already, the extrinsic theory involves us in nothing untoward, and simplifies and unifies our theory. We do not want mistakenly to impute causal powers to mathematical objects by using the extrinsic mathematical theory within physics. But, merely noting that mathematical objects are non-spatio-temporal blocks any such confusion.

Thus, the strength of the magic argument depends on whether we have a prior commitment to mathematical objects. The nominalist sees the uses of mathematics in science as magical, since s/he denies the existence of mathematical objects. The platonist sees no magic, only a demand for an account of the application of mathematics.

Several accounts of the application of mathematics in science, compatible with platonism, have been developed since Field's original monograph. Mark Balaguer argues that there is nothing magical about the utility of mathematics in physical science, since mathematics provides a theoretical apparatus for all possible physical states of affairs.[18] Pincock 2004 provides an explanation of the applications of mathematics which, though ontologically neutral, is compatible with platonism and so can also undermine Field's magic argument. Furthermore, one could deny the magic argument by

[18] See Balaguer 1998: 143.

claiming that mathematical objects are intrinsic to every physical system. Indeed, Quine's holism entails that the indispensabilist and the dispensabilist are committed to such a position, as I will discuss in the next section.

§7: PIE Does Not Favor Fictionalism

I have presented considerations which favor extrinsic theories over intrinsic ones, and which undermine PIE, and thus deflect Field's criticism of QI. In addition to Field's negative argument against QI4, he presents a positive account of mathematics, which he calls fictionalism. According to fictionalism, mathematical existence claims are false, and mathematical conditionals are vacuously true, if true. In this section, I argue that even if we accept PIE, it does not favor fictionalism. The indispensabilist can accept PIE since s/he should deny that mathematical objects are extrinsic to physical theory.

It is well-known that the indispensabilist has trouble accepting a distinction between abstract and concrete objects. As Charles Parsons notes,

> Although Quine makes some use of very general divisions among objects, such as between 'abstract' and 'concrete', these divisions do not amount to any division of *senses* either of the quantifier or the word 'object'; the latter sort of division would indeed call for a many-sorted quantificational logic rather than the standard one. Moreover, Quine does not distinguish between objects and any more general or different category of 'entities' (such as Frege's *functions*) (Parsons 1983: 377).

Furthermore, Quine himself wonders if such distinctions are sustainable.

> [O]dd findings [in quantum mechanics] suggest that the notion of a particle was only a rough conceptual aid, and that nature is better conceived as a distribution of local states over space-time. The points of space-time may be taken as quadruples of numbers, relative to some system of coordinates... We are down to an ontology of pure sets. The state functors remain as irreducibly physical vocabulary, but their arguments and values are pure sets. The ontological contrast between mathematics and nature lapses (Quine 1986: 402; see also Quine 1978; Quine 1960: 234; Quine 1974: 88; and Quine 1969: 98).

The indispensabilist's theory is constructed to explain or represent phenomena involving ordinary objects. "Bodies are assumed, yes; they are the things, first and foremost. Beyond them there is a succession of dwindling analogies" (Quine 1981: 9).

As these analogies dwindle, the traditional abstract/concrete distinction blurs, and so does the intrinsic/extrinsic distinction. For Quine, these distinctions must be made within science. But Quine's preferred theory does not support them. All of the indispensabilist's objects are posits of the same monolithic theory, made in the same way, for the same purposes of explaining our sense experience. There is no basis for discrete differences in type, no basis for either an intrinsic/extrinsic distinction or a related abstract/concrete distinction. Call this facet of indispensabilism ontic blur.

Field classified mathematical objects as extrinsic to NGT in part because of their causal isolation from physical ones. Field, though, is clearly thinking of traditional mathematical objects: abstract objects that exist in all possible worlds, and are knowable *a priori*, for example. Mark Balaguer defends a principle of causal isolation (PCI) governing the traditional separation of mathematical and physical objects. But, PCI is off limits to the indispensabilist. In fact, Balaguer notes that ontic blur is definitive of QI. "The Quine-Putnam argument should be construed as an argument not for platonism or the truth of mathematics but, rather, for the falsity of PCI" (Balaguer 1998: 110).

Just as the indispensabilist does not countenance traditional abstract objects, discretely distinct from ordinary objects, any dispensabilist must accept ontic blur. For, the dispensabilist accepts the indispensabilist's terms of debate, including QI1-3, which are the source of the blur. Given blur, Field can not call mathematical elements extrinsic to physical theory. The indispensabilist's mathematical objects are actually intrinsic to the one, holistic best theory.

Field tries to establish that the posits of space-time points differ from posits of mathematical objects in order to admit space-time points as intrinsic to physics. He claims that mathematical objects are supposed to be known *a priori*, while physical space is not (Field 1980: 31). But, for the indispensabilist, mathematical objects, like all objects, are known *a posteriori*. A defense of the apriority of our knowledge of mathematical objects would undermine both the dispensabilist and the indispensabilist, making Field's reformulation moot. An apriorist could argue for mathematical knowledge more forcefully, independently of QI.

Field also argues for the difference between the posits of mathematical objects and space-time points on the basis of the richer mathematical ideology (Field 1980: 32). But Resnik 1985 develops an impressive amount of mathematics within Field's space-time, the geometry of which corresponds to second-order analysis. Not only do we get addition and multiplication over the reals and the natural numbers, but we can set up a coordinate system, and define ordered n–tuples. We can even avoid the arbitrary choice of points to serve as 0 and 1 by substituting individual variables.

It is difficult to see how Field could deny that numbers are intrinsic to physical theories without turning PIE into some version of an eleatic principle, appealing to the causal isolation of mathematical objects. But, if he is presuming an eleatic principle, it is difficult to see why indispensability holds any sway. The eleatic can just deny Quine's argument in favor of a causal criterion for ontological commitment. The eleatic can be an instrumentalist about a theory's references, and need not reformulate physical theory to avoid commitments to mathematical objects.

Field's dispensabilist ideology and the indispensabilist's quasi-mathematical ideology both apply to intrinsic objects. The traditional platonist can make the extrinsic/intrinsic distinction. But by definition, the traditional platonist has an independent epistemology for mathematics. The dispensabilist reformulation of standard science does not denigrate our beliefs about mathematical objects if they are independently justified.

§8: Conclusion

It is difficult to see any value in PIE, as a general principle guiding theory choice. Resnik, reviewing Field's monograph, argues that we can see it at work in economics.

> The Expected Utility Theorem, which underwrites the use of utility functions, establishes that if an agent's preference ordering satisfies certain conditions then it can be represented by a real valued function which is unique up to positive linear transformations. From this it is usually argued that there is no need to presuppose ill understood utilities in accounting for behavior which maximizes expected utility because an account can be given directly in terms of preferences. (Resnik 1983: 515)

Resnik says that an intrinsic account, in terms of preferences, is desirable because utilities are ill understood. But, if they were better understood than preferences, then the account would go the other way. If we could order utilities uniquely, while remaining confused about inter- and intra-personal comparisons of preferences, we would seek to explain preferences in terms of utilities. One principle underlying Resnik's preference is that we should explain things we do not understand in terms of things we do understand. Appropriate Ockhamist principles also guide the avoidance of utilities. It is ironic that Resnik uses an example which employs mathematics to characterize the elements we understand. If utilities were as well understood as mathematical theories, then accounts in terms of them would be welcome. PIE is doing no work, here.

I have argued that our preference for unification of theories undermines PIE. A proponent of PIE might complain that once we introduce bridge principles which unify two distinct theories, they are no longer extrinsic to each other, and thus that PIE is not in conflict with unification. Before unification, we have separate theories, and explanations of the principles of one theory in terms of principles of the other would be disfavored. After unification, such explanations would be welcome.

To be slightly more precise, consider (the conjunction of axioms of) two completely independent theories, T_1 and T_2. We could take T_1 to be biology and T_2 to be quantum mechanics; or, we could take T_1 to be ZFC and T_2 to be general relativity. But, assume that T_1 and T_2 are indisputably extrinsic to each other. The theory $T_1 + T_2$ which merely conjoins two sets of axioms is thus an extrinsic theory. Explanation of phenomena governed by the axioms of T_1 in terms of the principles embodied in T_2 would be extrinsic explanations.

Now, consider a set of mapping principles, M, which bridge T_1 and T_2. We can see that Field thinks that $T_1 + T_2 + M$ is also an extrinsic theory by noting that standard (mathematized) physics includes physical axioms, mathematical axioms, and mappings between the two. These mapping principles are precisely at work when we measure the length of a wire in meters, or when we discuss the Hilbert space of an atom.

The defender of PIE who wishes to embrace unification claims that $T_1 + T_2 + M$ is an intrinsic theory, since the bridge principles connect the objects posited by T_1 with the the objects posited by T_2. This approach would save PIE. We could all agree that extrinsic explanations, in the sense of explanations that used $T_1 + T_2$ (without M), were magical, and to be disfavored. But, this interpretation of PIE deprives it of all application. For, on this view there would be no extrinsic explanations. We would never appeal to mathematics in physics, or to quantum mechanics in biology, unless we had bridge principles in hand. Any plausible explanation would have to be intrinsic. Unification really is opposed to intrinsic explanation.

We can appreciate both intrinsic and extrinsic theories. The situation is analogous to the relation between classical mathematicians and intuitionists, from a classical perspective. The classical mathematician can appreciate the distinction between constructive and nonconstructive proofs, without concluding that only constructive proofs tell us what exists. Similarly, we can appreciate the technical acuity of Field's construction without inferring that there are no mathematical objects.

Philosophers with nominalist predispositions may see PIE as a commonsense principle, and so may have neglected to recognize a gap in Field's argument against QI. There also may be other reasons to reject QI, or merely to prefer a theory which does not quantify over, or otherwise refer to, mathematical objects. But the principle of intrinsic explanations can not do this work.

References

Armstrong, David. 1978b. *A Theory of Universals: Universals and Scientific Realism, Volume II.* Cambridge: Cambridge University Press.

Armstrong, David. 1978a. *Nominalism and Realism: Universals and Scientific Realism, Volume I.* Cambridge: Cambridge University Press.

Azzouni, Jody. 2004. *Deflating Existential Consequence: A Case for Nominalism*. New York: Oxford University Press.

Baker, Alan. 2005. "Are there Genuine Mathematical Explanations of Physical Phenomena?" *Mind:* 114: 223-238.

Baker, Alan. 2003. "The Indispensability Argument and Multiple Foundations for Mathematics." *The Philosophical Quarterly* 53.210: 49-67.

Balaguer, Mark. 1998. *Platonism and Anti-Platonism in Mathematics*. New York: Oxford University Press.

Bangu, Sorin Ioan. 2008. "Inference to the Best Explanation and Mathematical Realism." *Synthese* 160: 13-20.

Burgess, John, and Gideon Rosen. 1997. *A Subject with No Object*. New York: Oxford.

Carnap, Rudolph. 1950. "Empiricism, Semantics, and Ontology." In his *Meaning and Necessity: A Study in Semantics and Modal Logic*, second edition. The University of Chicago Press, 1956; Midway reprint, 1988.

Colyvan, Mark. Forthcoming. "There's No Easy Road to Nominalism." *Mind*.

Colyvan, Mark. 2001. *The Indispensability of Mathematics*. Oxford University Press.

Colyvan, Mark. 1999. "Confirmation Theory and Indispensability." *Philosophical Studies* 96: 1-19.

Devitt, Michael. 1984 (1991, 1997). *Realism and Truth*, 2nd ed. with Afterword. Princeton: Princeton University Press.

Field, Hartry. 1990. "Mathematics Without Truth (A Reply to Maddy)." *Pacific Philosophical Quarterly* 71: 206-222.

Field, Hartry. 1989. *Realism, Mathematics, and Modality*. Oxford: Basil Blackwell.

Field, Hartry. 1985b. "On Conservativeness and Incompleteness." In Field 1989.

Field, Hartry. 1985a. "Can We Dispense with Space-Time?" Reprinted, with postscript, in Field 1989.

Field, Hartry. 1980. *Science Without Numbers*. Princeton: Princeton University Press.

Frege, Gottlob. 1980. *Philosophical and Mathematical Correspondence*. Chicago: University of Chicago Press.

Hilbert, David. 1971. *Foundations of Geometry*. La Salle: The Open Court Press.

Kitcher, Philip. 1981. “Explanatory Unification.” *Philosophy of Science* 48.4: 507-531

Leng, Mary. 2005. “Mathematical Explanation.” In Cellucci, Carlo and Donald Gillies eds, *Mathematical Reasoning and Heuristics*, King's College Publications, London, 167-189.

Lyon, Aidan and Mark Colyvan. 2007. “The Explanatory Power of Phase Spaces.” *Philosophia Mathematica* 16.2: 227-243.

Mancosu, Paolo. 2008. “Explanation in Mathematics.” *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/mathematics-explanation/>.

Melia, Joseph. 2000. “Weaseling Away the Indispensability Argument.” *Mind* 109: 455-479.

Oddie, Graham. 1982. “Armstrong on the Eleatic Principle and Abstract Entities.” *Philosophical Studies* 41: 285-95.

Parsons, Charles. 1983. “Quine on the Philosophy of Mathematics.” In Hahn, Lewis Edwin and Paul Arthur Schilpp (eds.). 1986. *The Philosophy of W.V. Quine*. La Salle: Open Court.

Pincock, Christopher. 2007. “A Role for Mathematics in the Physical Sciences.” *Nous* 41: 253-275.

Pincock, Christopher. 2004. “A New Perspective on the Problem of Applying Mathematics.” *Philosophia Mathematica* 12.2: 135-161.

Putnam, Hilary. 1981. “Convention: A Theme in Philosophy.” In his *Realism and Reason: Philosophical Papers, Vol. III*. Cambridge: Cambridge University Press.

Putnam, Hilary. 1971. *Philosophy of Logic*. Reprinted in his *Mathematics, Matter, and Method: Philosophical Papers, Vol. I*. Cambridge: Cambridge University Press, 1975.

Quine, W.V. 1986. “Reply to Charles Parsons.” In Hahn, Lewis Edwin and Paul Arthur Schilpp (eds.). 1986. *The Philosophy of W.V. Quine*. La Salle: Open Court.

Quine, W.V. 1981. “Things and Their Place in Theories.” In *Theories and Things*. Cambridge: Harvard University Press.

Quine, W.V. 1978. “Success and the Limits of Mathematization.” In *Theories and Things*. Cambridge: Harvard University Press.

Quine, W.V. 1974. *The Roots of Reference*. La Salle: The Open Court Press.

Quine, W.V. 1969. “Existence and Quantification.” In his *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Quine, W.V. 1960. *Word & Object*. Cambridge: The MIT Press.

Quine, W.V. 1958. “Speaking of Objects.” In *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Quine, W.V. 1955. "Posits and Reality." In his *The Ways of Paradox*. Cambridge: Harvard University Press, 1976.

Quine, W.V. 1951. "Two Dogmas of Empiricism." Reprinted in *From a Logical Point of View*. Cambridge: Harvard University Press, 1980.

Quine, W.V. 1948. "On What There Is." Reprinted in *From a Logical Point of View*. Cambridge: Harvard University Press, 1980.

Quine, W.V. 1939. "A Logistical Approach to the Ontological Problem." In *The Ways of Paradox*. Cambridge: Harvard University Press.

Resnik, Michael. 1997. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press.

Resnik, Michael D. 1985. "Ontology and Logic: Remarks on Hartry Field's Anti-platonist Philosophy of Mathematics." *History and Philosophy of Logic* 6: 191-209.

Resnik, Michael D. 1983. Review of Field 1980. *Nous*: 514-519.

Shapiro, Stewart. 1983. "Conservativeness and Incompleteness." *The Journal of Philosophy* 80.9: 521-531.

Facta Philosophica 9, 2007: 203–211

Structuralism, Indispensability, and the Access Problem

Russell Marcus

§1: Structuralism and the Problem of Mathematics

The central epistemological problem for mathematics is sometimes called the access problem. The access problem arises from a two-realm view, the supposition that the referents of mathematical singular terms inhabit a realm which is separate from us. The history of philosophy is littered with attempts to solve the access problem, which is just the old question of how concrete humans can have knowledge of abstract mathematical objects. Quine's approach in the philosophy of mathematics, his indispensability argument, dissolves the access problem, though his solution sometimes goes unrecognized, even by those who are sympathetic to his work. This paper shows that the proponent of the indispensability argument can, to a certain extent, deflect criticisms based on the access argument.

One way to try to solve the access problem involves positing a special human ability to learn about abstract objects. This rationalist approach is exemplified by Plato, Descartes, Gödel, and Jerrold Katz, among many others. The central complaints about rationalism involve accusations of mysticism and desperation, and failure of parsimony, both epistemological parsimony and parsimony of the resulting ontology.[1]

A second approach to the access problem denies the supposition that mathematical objects are separate from us. This approach is exemplified by Aristotle and Mill, and by Penelope Maddy's early work. The central complaints about the Aristotelian approach involve the insufficiency of concrete objects as models of mathematical theories.[2]

1 Gödel 1961 discusses a special faculty of mathematical insight. See Katz 1998 for a recent defense of rationalism. Field 1982 accuses the rationalist of mysticism and desperation; §B of Field 1989a is in the same spirit.

2 Maddy 1990 contains her early position; Balaguer 1994 criticizes that position.

Even more radically, one could deny, as Berkeley and Hartry Field do, that we have mathematical knowledge. Calling his position fictionalism, Field agrees that we have logical knowledge, but he denies that there are any mathematical objects to which we need access.[3] Against fictionalism, we seem to have mathematical knowledge, and we can not, with the fictionalist, just wish away the problem.

Quine's approach replaces questions of access with justifications of theories and their corresponding posits. As an example of how Quine's solution may be neglected, consider MacBride 2004's criticism of Michael Resnik's structuralist attempt to solve the access problem.[4] Resnik 1997 argues that our experience with concrete templates leads to knowledge of abstract patterns. He describes how our knowledge of patterns could have been posited by our intellectual ancestors in response to technological and intellectual needs. Resnik's account is a rational reconstruction of our ancestors' postulation of mathematical objects.

MacBride reasonably interprets Resnik's story as an attempt to solve the access problem. I will argue that MacBride is correct that Resnik's quasi-historical account is unhelpful in justifying our knowledge of mathematics. But Resnik's structuralism, depending on an indispensability argument, actually has no access problem. Such criticisms of Quine, and Quineans, based on access problems, are misdirected. I start with a discussion of Quine's argument, and then show that Resnik has no access problem to the objects of his mathematical ontology. At the end of the paper, I highlight a real problem with Quine's approach, which is deeper than either MacBride or Resnik indicates.

§2: Descriptive and Normative Epistemologies

There are at least two kinds of stories that one can tell which may be relevant to mathematical epistemology, indeed any epistemology. The first kind of story describes how we actually acquire our beliefs. The second kind of story attempts to justify or legitimate those beliefs. The

3 Field 1980 and Field 1982 discuss fictionalism. Field 1998 argues that we can account for the apparent objectivity of mathematics on the basis of the objectivity of logic.

4 Structuralism in the philosophy of mathematics is rooted in the argument of Benacerraf 1965 that the axioms of set theory are satisfiable by no unique set of objects. Different sets of objects, all in the same pattern, can satisfy the axioms. Thus, what the axioms define is really a pattern, or a structure.

second kind of story is prescriptive, or normative.[5] Why do I believe what I do, and should I?

Resnik's quasi-historical account of the acquisitions of mathematical beliefs describes how our ancestors might have come to posit abstract objects in response to practical problems. They might have recognized the advantages of idealization, and of recognizing commonalities.

> I hypothesize that using concretely written diagrams to represent and design patterned objects, such as temples, bounded fields, and carts, eventually led our mathematical ancestors to posit geometric objects as *sui generis*. With this giant step behind them it was and has been relatively easy for subsequent mathematicians to enlarge and enrich the structures they knew, and to postulate entirely new ones (Resnik 1997: 5).

Resnik's account is quasi-historical, and not really historical, since it is merely an attempt to describe the ways in which our ancestors could have come to believe in mathematical objects, and not an attempt to describe the ways in which they actually did come to believe in them. Resnik calls his approach postulational, and it is so in two ways. He makes ontological posits, of abstract objects. More relevantly here, Resnik makes epistemological posits of how our ancestors could have come to make those ontological posits. Resnik sees his task as defending these quasi-historical posits by their utility.

Resnik's quasi-historical account describes the ways in which one might come to acquire mathematical beliefs. It is not a prescriptive account. Exactly the same kind of story could be told about the utility of beliefs in Santa Claus or the Olympic gods, but we would not take these stories as justifying beliefs in Santa Claus or Zeus.

Evaluated on its descriptive merits alone, Resnik's quasi-historical story is quite odd. We do not really learn mathematics in history classes, by examining the thoughts or beliefs of those who come before us. Even if we tried to learn what our ancestors believed, we lack sufficient historical evidence of the beliefs of those who originally posited abstracta. Instead, a good descriptive account would rely on stories about our experiences in mathematics classes, or alone in studies with textbooks, failed attempts to solve problems, and late-night epiphanies.

Descriptive stories need not be irrelevant to justification. If you ask how I know that the man over there has three coins in his pocket, I might describe how I saw him place those coins in his pocket. Still, you may ask

5 A third kind of epistemic story involves questions of our ability to know whether claims are true. A fourth might refer to any other conditions we might place on a belief in order to consider it knowledge. These two further kinds of accounts need not concern us here.

further questions, not merely about chocolate coins and holes in pockets (i. e. about whether your perceptions are accurate in this particular case) but about the reliability of sense experience generally, and whether I know I am not now in a dream state. These latter questions concern whether reference to the genesis of a belief can suffice to justify that belief. They are prescriptive questions, and should lead us to the conclusion that descriptive stories are never sufficient for justification, by themselves. To justify any belief, I must supplement the description of how I acquire that belief with an account of why the belief-forming processes to which my descriptive account refers are in fact reliable, or conducive to truth, or whatever standard you may prefer. The insufficiency of descriptive stories is not limited to sense perception. If I believe a mathematical theorem on the basis of an a priori proof, I still must account for the reliability of the proof procedure. Indeed, taking a descriptive story as a sufficient account of knowledge is a kind of genetic fallacy.

Quine, as is well known, urged us to relegate epistemology, naturalized epistemology, to the description of how stimulus leads to science. Epistemology, Quine urges, should be a form of empirical psychology. Jaegwon Kim criticized Quine for eliminating epistemology's normative element. The naturalized epistemologist has a story to tell, but if that story omits a prescriptive justificatory account, it will have to be incomplete. "For epistemology to go out of the business of justification is for it to go out of business" (Kim 1988: 393).

Kim, I think, is correct that we need a normative account, but wrong that Quine fails to provide one. To see that Quine at least implicitly presents a normative epistemology, consider the following well-known facts. First, Quine preferred, where possible, desert landscapes; he would have liked to do without abstract objects if he could. Second, he did not think that he could, in fact, do without mathematical objects and so he admitted extensional ones, i.e. sets, on the basis of what has become known as the indispensability argument.

§3: Indispensability and Access

The indispensability argument says that our knowledge of the abstract objects of mathematics is justified by their ineliminable uses in empirical science.[6]

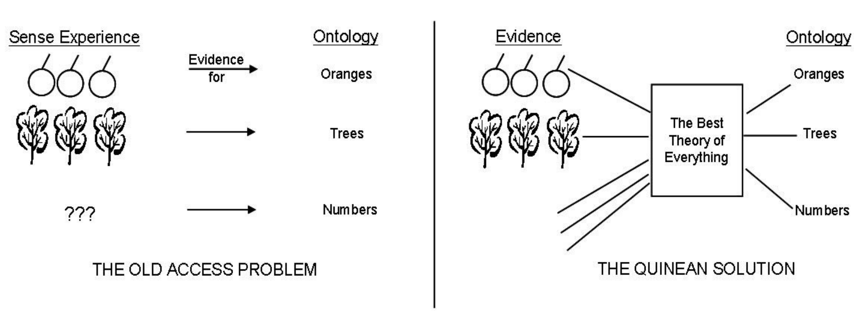
6 Quine nowhere presents a detailed indispensability argument, though he alludes to one in many places. Among them: Quine 1939, 1948, 1951, 1955, 1958, 1960, 1978, and 1986.

(QI) QI.1: We should believe the (single, holistic) theory which best accounts for our sense experience.
QI.2: If we believe a theory, we must believe in its ontic commitments.
QI.3: The ontic commitments of any theory are the objects over which that theory first-order quantifies.
QI.4: The theory which best accounts for our sense experience first-order quantifies over mathematical objects.
QI.C: We should believe that mathematical objects exist.

For the indispensabilist, no independent mathematical epistemology, beyond that for empirical science, is necessary. The indispensability argument dissolves the traditional access problem. According to the confirmation holism which underwrites Quine's indispensability argument, evidence for any claim is spread throughout the entirety of the theory which yields that claim. The indispensabilist needs no account of access to his so-called mathematical objects, since they are posited by a theory which takes only our sense experience as evidence. Even for ordinary objects the question of access is moot. All objects are posits.

To demand access is to demand that a perceiver be able to correlate the objects he or she believes exist with particular perceptions. The traditional empiricist requires lines of access from, say, the tree to my eyes, to my brain, to my beliefs. He or she draws roughly parallel lines to account for beliefs about all objects. The sense-data reductionist demands, as did Hume, a connection to sense experience for every legitimate claim.

Quine denies that a satisfactory account of piecemeal access, like that of the sense-data reductionist, is available. His methodology is a response to the difficulties of describing our access to all objects. Instead, Quine isolates evidence, on one side, and ontology on the other. Between them stands a theory which, as a whole, must be consistent with the evidence. The indispensability argument is merely a corollary of this procedure.



The indispensabilist argues that we posit sets as values of variables bound by the quantifiers of NF, for example, not to account for mathematical facts such as '2 + 2 = 4'. We posit them because we need them to account for the stimulations of our nerve endings by ordinary objects. Instead of an access problem, the indispensabilist has a problem of sufficiency of evidence for the theory.

One might wonder, still, given the two-realm view, if the indispensabilist has solved the access problem. In fact, implicitly, the indispensabilist has discarded the mathematical realm. All posits are equal, in the relevant sense. The mathematical objects in the indispensabilist's ontology need not be deemed apart from us. Since the indispensabilist's so-called mathematical objects are posits of a theory designed to account for our knowledge of ordinary, concrete objects, there is really no need to consider them as abstract. Indeed, the terms 'abstract' and 'concrete' become rather meaningless for the holist, vulgar terms in which the learned may only lightly indulge.

> In the case of abstract entities, certain protests against Platonism become irrelevant. There is no mysterious 'realm' of, say, sets in the sense that they need to have anything akin to location, and our knowledge of them is not based on any mysterious kind of 'seeing' into such a realm. This 'demythologizing' of the existence of abstract entities is one of Quine's important contributions to philosophy… (Parsons 1986: 377–78)

Though he is critical of Quine's method, Stewart Shapiro also relies on indispensability's dissolution of the mathematical realm. "My view is that, extensionally speaking, there is no…philosophically illuminating difference [between mathematical structures and other kinds of structures]." (Shapiro 1983: 542)

One might think that the fact that the indispensability argument dissolves rather than solves the access problem is a problem for the indispensabilist, and I do think this. But others, e. g. Resnik, take Quine's method as something like cutting the Gordian knot. We can, they say, have an abstract ontology with only an empiricist epistemology.

§4: Structuralism and Indispensability

Which brings us back to Resnik and MacBride. Resnik's structuralism relies on Quine's indispensability argument, and he makes no secret of it. In fact, the first third or so of Resnik 1997 is dedicated to a defense of the indispensability argument, which provides Resnik a normative epistemic story. Resnik appeals to the indispensability argument, and its holism, to dissolve the access problem for mathematical objects. The indispensabilist need not describe any access to any particular objects, including those of the purported second (abstract) world.

It is difficult to understand why Resnik presents his quasi-historical account in connection with our posits of structures, since it answers neither the normative nor the descriptive questions. The indispensability argument alone suffices to justify knowledge of its posits, so we do not need the quasi-historical story to provide the normative account. Further, we can not use it to answer the descriptive question because Resnik's postulation inaccurately describes our actual belief-forming processes. Perhaps Resnik thought that the indispensabilist's dissolution is inadequate. It easy to see why MacBride would make the mistake of thinking that the structuralist needs a solution to the access problem, since Resnik seems to think so too.

§5: A Real Access Problem

Lastly, there is another way in which the structuralist does have an access problem. According to the indispensabilist, there is no abstract realm into which one must see in order to know about mathematical objects. Still, the indispensabilist claims justification of our knowledge of mathematical objects.[7] One might think that it is essential to mathematical objects that they have certain properties traditionally ascribed to them. Mathematical objects are traditionally understood to be abstract, to lack spatio-temporal position, to be knowable a priori, and to exist necessarily. If we require that mathematical objects have these traditional properties, then the one-world picture is unsatisfactory. For, none of these properties follow on the indispensabilist's argument.

For example, consider necessity. Mathematical objects are traditionally thought to exist necessarily, perhaps since their existence does not appear to be contingent on any features of the concrete world. While some philosophers, like Quine, have impugned all modalities, necessity has a stubborn resilience. Even if we are convinced that there are deep problems with modalities, my existence seems contingent in a way that the existence of mathematical objects does not.

In contrast, according to the indispensabilist, mathematical objects are posited to account for our experience of a contingent world. If the world were different, it would require different objects. Suppose, for example, that charge really is a continuous property of real particles in

7 Quine, for example, never presented the indispensability argument as dissolving the two-realm picture in the way that Parsons does.

this world. The indispensabilist alleges that the world thus contains continuous functions. Further, suppose that in a different world, or if this world were different, there were no continuous properties. In that alternative world, says the indispensabilist, there are no continuous functions. Similar arguments can easily be constructed for other traditional mathematical properties.

One might argue that the indispensabilist's so-called mathematical objects, lacking so many traditional properties, are not really mathematical objects. They are empirical posits, made to account for the stimulations of our senses, which are merely less tractable than ordinary objects. The indispensabilist's mathematical objects are to electrons, say, as electrons are to tables. Though these so-called mathematical objects are not themselves sensed, they are no different in kind from other posits. Thus, goes the objection, the structuralist who relies on an indispensability argument has no basis on which to claim knowledge of mathematical objects. This is a real access problem, and nothing in structuralism, or the indispensability argument, seems likely to solve, or dissolve, it.

Russell Marcus
Department of Philosophy
Hamilton College
Clinton NY 13323
USA
rmarcus1@hamilton.edu

References

Balaguer, Mark. 1994. "Against (Maddian) Naturalized Platonism." *Philosophia Mathematica (3)*, 2: 97–108.

Benacerraf, Paul. 1965. "What Numbers Could Not Be." In Benacerraf and Putnam, eds., *Philosophy of Mathematics: Selected Readings*, second edition. Cambridge: Cambridge University Press, 1983.

Field, Hartry. 1998. "Mathematical Objectivity and Mathematical Objects." In *Truth and the Absence of Fact*, Oxford: Clarendon Press, 2001.

Field, Hartry. 1989b. *Realism, Mathematics, and Modality*. Oxford: Basil Blackwell.

Field, Hartry. 1989a. "Fictionalism, Epistemology and Modality." Introduction to Field 1989b.

Field, Hartry. 1982. "Realism and Anti-realism about Mathematics." Reprinted, with postscript, in Field 1989b.

Gödel, Kurt. 1961. "The Modern Development of the Foundations of Mathematics in the Light of Philosophy." In *Kurt Gödel: Collected*

Works, Vol. III., Solomon Feferman, et al., eds. 1995: New York, Oxford University Press.

Katz, Jerrold J. 1998. *Realistic Rationalism*. Cambridge: The MIT Press.

Kim, J. 1988. "What is Naturalized Epistemology?" In James E. Tomberlin, ed, *Philosophical Perspectives 2*, Asascadero, CA: Ridgeview Publishing Co: 381–406.

MacBride, F. 2004. "Can Structuralism Solve the 'Access' Problem?" *Analysis* 64: 309–318.

Maddy, Penelope. 1990. *Realism in Mathematics*. Oxford: Clarendon Press.

Parsons, C. 1986. "Quine on the Philosophy of Mathematics." In *The Philosophy of W. V. Quine*, ed. L.E. Hahn and P.A. Schilpp. La Salle: Open Court.

Quine, W.V. 1986. "Reply to Charles Parsons." In *The Philosophy of W. V. Quine*, ed. L.E. Hahn and P.A. Schilpp. La Salle: Open Court.

Quine, W.V. 1980. *From a Logical Point of View*. Cambridge: Harvard University Press.

Quine, W.V. 1978. "Success and the Limits of Mathematization." In his *Theories and Things*. Cambridge: Harvard University Press, 1981.

Quine, W.V. 1971. "Epistemology Naturalized." In Quine (1969a).

Quine, W.V. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Quine, W.V. 1960. *Word & Object*. Cambridge: The MIT Press.

Quine, W.V. 1958. "Speaking of Objects." In Quine (1969a).

Quine, W.V. 1955. "Posits and Reality." In his *The Ways of Paradox*. Cambridge: Harvard University Press, 1976.

Quine, W.V. 1951. "Two Dogmas of Empiricism." In Quine (1980).

Quine, W.V. 1948. "On What There Is." In Quine (1980).

Quine, W.V. 1939. "A Logistical Approach to the Ontological Problem." In Quine (1976a).

Quine, W.V. and N. Goodman. 1947. "Steps Toward a Constructive Nominalism." In Goodman, Nelson. *Problems and Projects*. Indianapolis: The Bobbs-Merrill Company, Inc., 1972.

Resnik, M. 1997. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press.

Shapiro, S. 1983. "Mathematics and Reality." *Philosophy of Science* 50: 523–548.

The Indispensability Argument in the Philosophy of Mathematics

In *The Internet Encyclopedia of Philosophy*, First Posted: October 18, 2010.

Russell Marcus
Department of Philosophy, Hamilton College
198 College Hill Road
Clinton NY 13323
rmarcus1@hamilton.edu
(315) 859-4056 (office)
(315) 381-3125 (home)

§0. Introduction

In his seminal 1973 paper, "Mathematical Truth", Paul Benacerraf presents a problem facing all accounts of mathematical truth and knowledge. Standard readings of mathematical claims entail the existence of mathematical objects. But, our best epistemic theories seem to debar any knowledge of mathematical objects. Thus, the philosopher of mathematics faces a dilemma: either abandon standard readings of mathematical claims or give up our best epistemic theories. Neither option is attractive.

The indispensability argument in the philosophy of mathematics is an attempt to avoid Benacerraf's dilemma by showing that our best epistemology is consistent with standard readings of mathematical claims. Broadly speaking, it is an attempt to justify knowledge of an abstract mathematical ontology using only a strictly empiricist epistemology.

The indispensability argument in the philosophy of mathematics, in its most general form, consists of two premises. The major premise states that we should believe that mathematical objects exist if we need them in our best scientific theory. The minor premise claims that we do in fact require mathematical objects in our scientific theory. The argument concludes that we should believe in the abstract objects of mathematics.

This article begins with a general overview of the problem of justifying our mathematical beliefs that motivates the indispensability argument. The most prominent proponents of the indispensability argument have been W.V. Quine and Hilary Putnam. The second section of the article discusses a reconstruction of Quine's argument in detail. Quine's argument depends on his general procedure for determining a best theory and its ontic commitments, and on his confirmation holism. The third and fourth sections of the article discuss versions of the indispensability argument which do not depend on Quine's method: one from Putnam and one from Michael Resnik.

The relationship between constructing a best theory and producing scientific explanations has recently become a salient topic in discussions of the indispensability argument. The fifth section of the article discusses a newer version of the indispensability argument, the explanatory indispensability argument. The last four sections of the article are devoted to a general characterization of indispensability arguments over various version (Section 6); a brief discussion of the most prominent responses to the indispensability argument (Section 7); a distinction between inter- and intra-theoretic indispensability arguments (Section 8); and a short conclusion.

Table of Contents

1. The Problem of Beliefs About Mathematical Objects
2. Quine’s Indispensability Argument
    2a. A Best Theory
    2b. Believing Our Best Theory
    2c. Quine’s Procedure for Determining Ontic Commitments
    2d. Mathematization
3. Putnam’s Success Argument
4. Resnik’s Pragmatic Indispensability Argument
5. The Explanatory Indispensability Argument
6. Characteristics of Indispensability Arguments in the Philosophy of Mathematics
7. Responses to the Indispensability Argument
8. Inter-theoretic and Intra-theoretic Indispensability Arguments
9. Conclusion

§1. The Problem of Beliefs About Mathematical Objects

Most of us have a lot of mathematical beliefs. For example, one might believe that the tangent to a circle intersects the radius of that circle at right angles, that the square root of two can not be expressed as the ratio of two integers, and that the set of all subsets of a given set has more elements than the given set. These claims all refer to mathematical objects, including circles, integers, and sets. The fundamental question that motivates the indispensability argument is, "How can we justify our mathematical beliefs?"

Mathematical objects are in many ways unlike ordinary physical objects like trees and cars. We learn about ordinary objects, at least in part, by using our senses. It is not obvious that we learn about mathematical objects this way. Indeed, it is difficult to see how we could use our senses to learn about mathematical objects. We do not see integers, or hold sets. Even geometric figures are not the kinds of things that we can sense. Consider any point in space; call it P. P is only a point, too small for us to see, or otherwise sense. Now imagine a precise fixed distance away from P, say an inch and a half. The collection of all points that are exactly an inch and a half away from P is a sphere. The points on the sphere are, like P, too small to sense. We have no sense experience of the geometric sphere. If we tried to approximate the sphere with a physical object, say by holding up a ball with a three-inch diameter, some points on the edge of the ball would be slightly further than an inch and a half away from P, and some would be slightly closer. The sphere is a mathematically precise object. The ball is rough around the edges. In order to mark the differences between ordinary objects and mathematical objects, we often call mathematical objects abstract objects.

When we study geometry, the theorems we prove apply directly and exactly to mathematical objects, like our sphere, and only indirectly and approximately to physical objects, like our ball. Numbers, too, are insensible. While we might see or touch a bowl of precisely eighteen grapes, we see and taste the grapes, not the eighteen. We can see a numeral, '18', but that is the name for a number, just as the term 'Russell' is my name and not me. We can sense the elements of some sets, but not the sets themselves. And some sets are sets of sets, abstract collections of abstract objects. Mathematical objects are not the kinds of things that we can see or touch, or smell, taste or hear. If we can not learn about mathematical objects by using our senses, a serious worry arises about how we can justify our mathematical beliefs.

The question of how we can justify our beliefs about mathematical objects has long puzzled philosophers. One obvious way to try to answer our question is to appeal to the fact that we prove the theorems of mathematics. But, appealing to mathematical proofs will not solve the problem. Mathematical proofs are ordinarily construed as derivations from fundamental axioms. These axioms, e.g. the Zermelo-Frankel axioms for set theory, the Peano Axioms for number theory, or the more-familiar Euclidean axioms for geometry, refer to the same kinds of mathematical objects. Our question remains about how we justify our beliefs in the axioms.

For simplicity, to consider our question of how we can justify our beliefs about mathematical objects, we will only consider sets. Set theory is generally considered the most fundamental mathematical theory. All statements of number theory, including those concerning real numbers, can be written in the language of set theory. Through the method of analysis, all geometric theorems can be written as algebraic statements, where geometric points are represented by real numbers. Claims from all other areas of mathematics can be written in the language of set theory, too.

Sets are abstract objects, lacking any spatio-temporal location. Their existence is not contingent on our existence. They lack causal efficacy. Our question, then, given that we lack sense experience of sets, is how we can justify our beliefs about sets and set theory.

There are a variety of distinct answers to our question. Some philosophers, called rationalists, claim that we have a special, non-sensory capacity for understanding mathematical truths, a rational insight arising from pure thought. But, the rationalist's claims appear incompatible with an understanding of human beings as physical creatures whose capacities for learning are exhausted by our physical bodies. Other philosophers, called logicists, argue that mathematical truths are just complex logical truths. In the late nineteenth and early twentieth centuries, the logicists Gottlob Frege, Alfred North Whitehead, and Bertrand Russell attempted to reduce all of mathematics to obvious statements of logic, e.g. that every object is identical to itself, or that if p then p. But, it turns out that we can not reduce mathematics to logic without adding substantial portions of set theory to our logic. A third group of philosophers, called nominalists or fictionalists, deny that there are any mathematical objects; if there are no mathematical objects, we need not justify our beliefs about them.

The indispensability argument in the philosophy of mathematics is an attempt to justify our mathematical beliefs about abstract objects, while avoiding any appeal to rational insight. Its most significant proponent was Willard van Orman Quine.

§2. Quine's Indispensability Argument

Though Quine alludes to an indispensability argument in many places, he never presented a detailed formulation. For a selection of such allusions, see Quines 1939, 1948, 1951, 1955, 1958, 1960, 1978, and 1986a. I will discuss the following version of the argument:

QI
- QI1. We should believe the theory which best accounts for our sense experience.
- QI2. If we believe a theory, we must believe in its ontic commitments.
- QI3. The ontic commitments of any theory are the objects over which that theory first-order quantifies.
- QI4. The theory which best accounts for our sense experience first-order quantifies over mathematical objects.
- QIC. We should believe that mathematical objects exist.

One presumption behind QI is that the theory which best accounts for our sense experience is our best scientific theory. Thus, Quine naturally defers much of the work of determining what exists to scientists. While it is obvious that scientists use mathematics in developing their theories, it is not obvious why the uses of mathematics in science should lead us to believe in the existence of abstract objects. For example, when we study the interactions of charged particles, we rely on Coulomb's Law, which states that the electromagnetic force between two charged particles is proportional to the charges on the particles and, inversely, to the distance between them.

CL $\quad F = k\,|q_1 q_2| / r^2$, where the electrostatic constant $k \approx 9 \times 10^9\ Nm^2/c^2$

CL refers to a real number, k, and employs mathematical functions like multiplication and absolute value. Still, we use Coulomb's Law to study charged particles, not to study mathematical objects, which have no effect on those particles. The plausibility of Quine's indispensability argument thus depends on both Quine's claim that the evidence for our scientific theories transfers to the mathematical elements of those theories, which is implicit in QI1, and his method for determining the ontic commitments of our theories at QI3 and QI4. The method underlying Quine's argument involves gathering our physical laws and writing them in a canonical language of first-order logic. The commitments of this formal theory may be found by examining its quantifications.

The remainder of this section discusses each of the premises of QI in turn.

§2a: A Best Theory

The first premise of QI is that we should believe the theory which best accounts for our sense experience, i.e. we should believe our best scientific theory.

Quine's belief that we should defer all questions about what exists to natural science is really an expression of what he calls, and has come to be known as, naturalism. He describes naturalism as, "[A]bandonment of the goal of a first philosophy. It sees natural science as an inquiry into reality, fallible and corrigible but not answerable to any supra-scientific tribunal, and not in need of any justification beyond observation and the hypothetico-deductive method" (Quine 1981: 72).

Quine developed his naturalism in large part as a response to logical positivism, which is also called logical empiricism or just positivism. Positivism requires that any justified belief be constructed out of, and be reducible to, claims about observable phenomena. We know about ordinary objects like trees because we have sense experience, or sense data, of trees directly. We know about very small or very distant objects, despite having no direct sense experience of them, by having sense data of their effects, say electron trails in a cloud chamber. For the positivists, any scientific claim must be reducible to sense data.

Instead of starting with sense data and reconstructing a world of trees and persons, Quine assumes that ordinary objects exist. Further, Quine starts with an understanding of natural science as our best account of the sense experience which gives us beliefs about ordinary objects. Traditionally, philosophers believed that it was the job of epistemology to justify our knowledge. In contrast, the central job of Quine's naturalist is to describe how we construct our best theory, to trace the path from stimulus to science, rather than to justify knowledge of either ordinary objects or scientific theory.

Quine's rejection of positivism included the insight, now known as confirmation holism, that individual sentences are only confirmed in the context of a broader theory. Confirmation holism arises from an uncontroversial observation that any sentence s can be assimilated without contradiction to any theory T, as long as we adjust truth values of any sentences of T that conflict with s. These adjustments may entail further adjustments, and the new theory may in the end look quite different than it did before we accommodated s. But we can, as a matter of logic, hold on to any sentence come what may. And, we are not forced to hold on to any statement, come what may; there are no unassailable truths.

For a simple example, suppose I have a friend named Abigail. I have a set of beliefs, which we can call a theory of my friendship with her. (A theory is just a collection of sentences.) Suppose that I overhear Abigail saying mean things about me. New evidence conflicts with my old theory. I have a choice whether to reject the evidence (e.g. "I must have mis-heard") or to accommodate the evidence by adjusting my theory, giving up the portions about Abigail being my friend, say, or about my friends not saying mean things about me. Similarly, when new astronomical evidence in the 15th and 16th centuries threatened the old geocentric model of the universe, people were faced with a choice of whether to accept the evidence, giving up beliefs about the Earth being at the center of the universe, or to reject the evidence. Instead of requiring that individual experiences are each independently assessed, confirmation holism entails that there are no justifications for particular claims independent of our entire best theory. We always have various options for restoring an inconsistent theory to consistency.

Confirmation holism entails that our mathematical theories and our scientific theories are linked, and that our justifications for believing in science and mathematics are not independent. When new evidence conflicts with our current scientific theory, we can choose to adjust either scientific principles or mathematical ones. Evidence for the scientific theory is also evidence for the mathematics used in that theory.

The question of how we justify our beliefs about mathematical objects arose mainly because we could not perceive them directly. By rejecting positivism's requirement for reductions of scientific claims to sense data, Quine allows for beliefs in mathematical objects despite their abstractness. We do not need sensory experience of mathematical objects in order to justify our mathematical beliefs. We just need to show that mathematical objects are indispensable to our best theory.

QI1 may best be seen as a working hypothesis in the spirit of Ockham's razor. We look to our most reliable endeavor, natural science, to tell us what there is. We bring to science a preference that it account for our entrenched esteem for ordinary experience. And we posit no more than is necessary for our best scientific theory.

§2b: Believing Our Best Theory

The second premise of QI states that our belief in a theory naturally extends to the objects which that theory posits.

Against QI.2, one might think that we could believe a theory while remaining agnostic or instrumentalist about whether its objects exist. Physics is full of fictional idealizations, like infinitely long wires, centers of mass, and uniform distributions of charge. Other sciences also posit objects that we do not really think exist, like populations in Hardy-Weinberg equilibrium (biology), perfectly rational consumers (economics), and average families (sociology). We posit such ideal objects to facilitate using a theory. We might believe our theory while recognizing that the objects to which it refers, strictly speaking, are only ideal. If we hold this instrumentalist attitude toward average families and infinitely long wires, we might hold it toward circles, numbers and sets, too.

Quine argues that any discrepancy between our belief in a theory and our beliefs in its objects is illegitimate double-talk. One can not believe in only certain elements of a theory which one accepts. If we believe a theory which says that there are centers of mass, then we are committed to those centers of mass. If we believe a theory which says that there are electrons and quarks and other particles too small to see, then we are committed to such particles. If our best theory posits mathematical objects, then we must believe that they exist.

QI1 and QI2 together entail that we should believe in all of the objects that our best theory says exist. Any particular evidence applies to the whole theory, which produces its posits uniformly. Quine thus makes no distinction between justifications of observable and unobservable objects, or between physical and mathematical objects. All objects, trees and electrons and sets, are equally posits of our best theory, to be taken equally seriously. "To call a posit a posit is not to patronize it" (Quine 1960: 22).

There will be conflict between our currently best theory and the better theories that future science will produce. The best theories are, of course, not now available. Yet, what exists does not vary with our best theory. Thus, any current expression of our commitments is at best speculative. We must have some skepticism toward our currently best theory, if only due to an inductive awareness of the transience of such theories.

On the other hand, given Quine's naturalism, we have no better theory from which to evaluate the posits of our currently best theory. There is no external, meta-scientific perspective. We know, casually and meta-theoretically, that our current theory will be superceded, and that we will give up some of our current beliefs. But, we do not know how our theory will be improved, and we do not know which beliefs we will give up. The best we can do is believe the best theory we have, and believe in its posits, and have a bit of humility about these beliefs.

§2c: Quine's Procedure for Determining Ontic Commitments

The first two premises of QI entail that we should believe in the posits of our best theory, but they do not specify how to determine precisely what the posits of a theory are. The third premise appeals to Quine's general procedure for determining the ontic commitments of a theory. Anyone who wishes to know what to believe exists, in particular whether to believe that mathematical objects exist, needs a general method for determining ontic commitment. Rather than relying on brute observations, Quine provides a simple, broadly applicable procedure. First, we choose a best theory. Next, we regiment that theory in first-order logic with identity. Last, we examine the domain of quantification of the theory to see what objects the theory needs to come out as true.

Quine's method for determining our commitments applies to any theory. Theories which refer to trees, electrons, and numbers, and theories which refer to ghosts, caloric, and God, are equally amenable of Quine's general procedure.

We have already discussed how the first step in Quine's procedure applies to QI. The second step of Quine's procedure for determining the commitments of a theory refers to first-order logic as a canonical language. Quine credits first-order logic with extensionality, efficiency, elegance, convenience, simplicity, and beauty. (See Quine 1986: 79 and 87.) Quine's enthusiasm for first-order logic largely derives from various attractive technical virtues. In first-order logic, a variety of definitions of logical truth concur: in terms of logical structure, substitution of sentences or of terms, satisfaction by models, and proof. First-order logic is complete, in the sense that any valid formula is provable. Every consistent first-order theory has a model. First-order logic is compact, which means that any set of first-order axioms will be consistent if every finite subset of that set is consistent. It admits of both upward and downward Löwenheim-Skolem theorems, which mean that every theory which has an infinite model will have a model of every infinite cardinality (upward) and that every theory which has an infinite model of any cardinality will have a denumerable model (downward). (See Mendelson 1997: 377.)

Less technically, the existential quantifier in first-order logic is a natural cognate of the English term 'there is', and Quine proposes that all existence claims can and should be made by existential sentences of first-order logic. "The doctrine is that all traits of reality worthy of the name can be set down in an idiom of this austere form if in any idiom" (Quine 1960: 228).

We should take first-order logic as our canonical language only if:

A. We need a single canonical language;
B. It really is adequate; and
C. There is no other adequate language.

Condition A arises almost without argument from QI1 and QI2. One of Quine's most striking and important innovations was his linking of our questions about what exists with our concerns when constructing a canonical formal language. When we regiment our correct scientific theory correctly, Quine argues, we will know what we should believe exists. "The quest of a simplest, clearest overall pattern of canonical notation is not to be distinguished from a quest of ultimate categories, a limning of the most general traits of reality" (Quine 1960: 161).

Whether condition B holds depends on how we use our canonical language. First-order logic is uncontroversially useful for what Quine calls semantic ascent. When we ascend, we talk about words without presuming that they refer to anything; we can deny the existence of objects without seeming to commit to them. For example, on some theories of language sentences which contain terms that do not refer to real things are puzzling. Consider:

CP The current president of the United States does not have three children.
TF The tooth fairy does not exist.

If CP is to be analyzed as saying that there is a current president who lacks the property of having three children, then by parity of reasoning TF seems to say that there is a tooth fairy that lacks the attribute of existence. This analysis comes close to interpreting the reasonable sentence TF as a contradiction saying that there is something that is not.

In contrast, we can semantically ascend, claiming that the term 'the tooth fairy' does not refer. TF is conveniently regimented in first-order logic, using 'T' to stand for the property of being the tooth fairy. '$\sim(\exists x)Tx$' carries with it no implication that the tooth fairy exists. Similar methods can be applied to more serious existence questions, like whether there is dark energy or an even number greater than two which is not the sum of two primes. Thus, first-order logic provides a framework for settling disagreements over existence claims.

Against the supposed adequacy of first-order logic, there are expressions whose first-order logical regimentations are awkward at best. Leibniz's law, that identical objects share all properties, seems to require a second-order formulation. Similarly, H resists first-order logical treatment.

H Some book by every author is referred to in some essay by every critic (Hintikka 1973: 345).

H may be adequately handled by branching quantifiers, which are not elements of first-order logic.

There are other limitations on first-order logic. Regimenting a truth predicate in first-order logic leads naturally to the paradox of the liar. And propositional attitudes, like belief, create opaque contexts that prevent natural substitutions of identicals otherwise permitted by standard first-order inference rules. Still, defenders of first-order logic have proposed a variety of solutions to these difficulties, many of which may not be due to first-order logic itself, but to deeper problems with language.

For condition C, Quine argues that no other language is adequate for canonical purposes. Ordinary language appears to be too sloppy, in large part due to its use of names to refer to objects. We use names to refer to some of the things that exist: 'Muhammad Ali', 'Jackie Chan', 'The Eiffel Tower'. But some names, like 'Spiderman', do not refer to anything real. Some things, like most insects and pebbles, lack names. Some things, like most people, have multiple names. We could clean up our language, constructing an artificial version in which everything has exactly one name. Still, in principle, there will not be enough names for all objects. As Cantor's diagonal argument shows, there are more real numbers than there are available names for those numbers. If we want a language in which to express all and only our commitments, we have to look beyond languages which rely on names.

Given the deficiencies of languages with names and Quine's argument that the existential quantifier is a natural formal equivalent of 'there is', the only obvious remaining alternatives to first-order logic as a canonical language are higher-order logics. Higher-order logics have all of the expressive powers of first-order logic, and more. Most distinctly, where first-order logic allows variables only in the object position (i.e. following a predicate), second-order logic allows variables in the predicate positions as well, and introduces quantifiers to bind those predicates. Logics of third and higher order allow further predication and quantification. As they raise no significant philosophical worries beyond those concerning second-order logic, this discussion will focus solely on second-order logic.

To see how second-order logic works, consider the inference R.

R	R1. There is a red shirt.
	R2. There is a red hat.
	RC. So, there is something (redness) that some shirt and some hat share.

RC does not follow from R1 and R2 in first-order logic, but it does follow in second-order logic.

R_S:	$R1_S$. $(\exists x)(Sx \bullet Rx)$
	$R2_S$. $(\exists x)(Hx \bullet Rx)$
	RC_S. $(\exists P)(\exists x)(\exists y)(Sx \bullet Hy \bullet Px \bullet Py)$

Accommodating inferences such as R by extending one's logic might seem useful. But, higher-order logics allow us to infer, as a matter of logic, that there is some thing, presumably the property of redness, that the shirt and the hat share. It is simple common sense that shirts and hats exist. It is a matter of significant philosophical controversy whether properties like redness exist. Thus, a logic which permits an inference like R_S is controversial.

Quine's objection to higher-order logics, and thus his defense of first-order logic, is that we are forced to admit controversial elements as interpretations of predicate variables. Even if we interpret predicate variables in the least controversial way, as sets of objects that have those properties, higher-order logics demand sets. Thus, Quine calls second-order logic, "Set theory in sheep's clothing" (Quine 1986a: 66). Additionally, higher-order logics lack many of the technical virtues, like completeness and compactness, of first-order logic. (For a defense of second-order logic, see Shapiro 1991.)

Once we settle on first-order logic as a canonical language, we must specify a method for determining the commitments of a theory in that language. Reading existential claims seems straightforward. For example, R2 says that there is a thing which is a hat, and which is red. But, theories do not determine their own interpretations. Quine relies on standard Tarskian model-theoretic methods to interpret first-order theories. On a Tarskian interpretation, or semantics, we ascend to a metalanguage to construct a domain of quantification for a given theory. We consider whether sequences of objects in the domain, taken as values of the variables bound by the quantifiers, satisfy the theory's statements, or theorems. The objects in the domain that make the theory come out true are the commitments of the theory. "To be is to be a value of a variable" (Quine 1939: 50, and elsewhere; for an accessible discussion of Tarskian semantics, see Tarski 1944).

Quine's procedure for determining the commitments of a theory can prevent us from prejudging what exists. We construct scientific theory in the most effective and attractive way we can. We balance formal considerations, like the elegance of the mathematics involved, with an attempt to account for the broadest sensory evidence. The more comprehensive and elegant the theory, the more we are compelled to believe it, even if it tells us that the world is not the way we thought it is. If the theory yields a heliocentric model of the solar system, or the bending of rays of light, then we are committed to heliocentrism or bent light rays. Our commitments are the byproducts of this neutral process.

§2d: Mathematization

The final step of QI involves simply looking at the domain of the theory we have constructed. When we write our best theory in our first-order language, we discover that the theory includes physical laws which refer to functions, sets, and numbers. Consider again Coulomb's Law: $F = k\,|q_1q_2|/r^2$. Here is a partial first-order regimentation which suffices to demonstrate the commitments of the law, using 'Px' for 'x is a charged particle'.

CLR $\forall x\forall y\{(Px \wedge Py) \supset (\exists f)[<q(x), q(y), d(x,y), k, F> \mid F = (k \cdot |q(x) \cdot q(y)|) / d(x,y)^2\}$

In addition to the charged particles over which the universal quantifiers in front range, there is an existential quantification over a function, f. This function maps numbers (the Coulomb's Law constant, and measurements of charge and distance) to other numbers (measurements of force between the particles).

In order to ensure that there are enough sets to construct these numbers and functions, our ideal theory must include set-theoretic axioms, perhaps those of Zermelo-Fraenkel set theory, ZF. The full theory of ZF is unnecessary for scientific purposes; there will be some sets which are never needed, some numbers which fail to measure any real quantity. But, we take a full set theory in order to make our larger theory as elegant as possible. We can derive from the axioms of any adequate set theory a vast universe of sets. So, CL contains or entails several existential mathematical claims. According to QI, we should believe that these mathematical objects exist.

Examples like CLR abound. Real numbers are used for measurement throughout physics, and other sciences. Quantum mechanics makes essential use of Hilbert spaces and probability functions. The theory of relativity invokes the hyperbolic space of Lobachevskian geometry. Economics is full of analytic functions. Psychology uses a wide range of statistics.

Opponents of the indispensability argument have developed sophisticated strategies for re-interpreting apparently ineliminable uses of mathematics, especially in physics. Some reinterpretations use alethic modalities (necessity and possibility) to replace mathematical objects. Others replace numbers with space-time points or regions. It is quite easy, if technical, to rewrite first-order theories in order to avoid quantifying over mathematical objects. It is less easy to do so while maintaining Quine's canonical language of first-order logic.

For example, Hartry Field's reformulation of Newtonian gravitational theory (Field 1980; see below, §7) replaces the real numbers which are ordinarily used to measure fundamental properties like mass and momentum with relations among regions of space-time. Field replaces the '2' in claims like, "The beryllium sphere has a mass of 2 kg" with a ratio of space-time regions, one twice as long as the other. In order to construct the proper ratios of space-time regions, and having no mathematical axioms at his disposal, Field's project requires either second-order logic or axioms of mereology, both of which are controversial extensions of first-order logic. For an excellent survey of dispensabilist strategies, and further references, see Burgess and Rosen 1997; for more recent work, see Melia 1998 and Melia 2000.

Quine's indispensability argument depends on controversial claims about believing in a single best theory, finding our commitments by using a canonical first-order logic, and the ineliminability of mathematics from scientific theories. Other versions of the argument attempt to avoid some of these controversial claims.

§3. Putnam's success argument

In his early work, Hilary Putnam accepted Quine's version of the indispensability argument. But, he came to differ with Quine on a variety of questions. Most relevantly, Putnam abandoned Quine's commitment to a single, regimented, best theory and he urged that realism in mathematics can be justified by its indispensability for correspondence notions of truth (which require set-theoretic relations) and for formal logic, especially for metalogical notions like derivability and validity which are ordinarily treated set-theoretically.

The position Putnam calls realism in mathematics is ambiguous between two compatible views. Sentence realism is the claim that sentences of mathematics can be true or false. Object realism is the claim that mathematical objects exist. Most object realists are sentence realists, though some sentence realists, including some structuralists, deny object realism. Indispensability arguments may be taken to establish either sentence realism or object realism. Quine was an object realist. Michael Resnik presents an indispensability argument for sentence realism; see below, §4. I take Putnam's realism to be both object and sentence realism, but nothing I say below depends on that claim.

Realism contrasts most obviously with fictionalism, on which there are no mathematical objects, and many mathematical sentences considered to be true by the realist are taken to be false. To understand the contrast between realism and fictionalism, consider the following two paradigm mathematical claims, the first existential and the second conditional.

E There is a prime number greater than any you have ever thought of.
C The consecutive angles of any parallelogram are supplementary.

The fictionalist claims that mathematical existence claims like E are false since prime numbers are numbers and there are no numbers. The standard conditional interpretation of C is that if any two angles are consecutive in a parallelogram, then they are supplementary. If there are no mathematical objects, then standard truth-table semantics for the material conditional entail that C, having a false antecedent, is true. The fictionalist claims that conditional statements which refer to mathematical objects, like C, are only vacuously true, if true.

Putnam's non-Quinean indispensability argument, the success argument, is a defense of realism over fictionalism, and other anti-realist positions. The success argument emphasizes the success of science, rather than the construction and interpretation of a best theory.

PS
PS1. Mathematics succeeds as the language of science.
PS2. There must be a reason for the success of mathematics as the language of science.
PS3. No positions other than realism in mathematics provide a reason.
PSC. So, realism in mathematics must be correct.

Putnam's success argument for mathematics is analogous to his success argument for scientific realism. The scientific success argument relies on the claim that any position other than realism makes the success of science miraculous. The mathematical success argument claims that the success of mathematics can only be explained by a realist attitude toward its theorems and objects. “I believe that the positive argument for realism [in science] has an analogue in the case of mathematical realism. Here too, I believe, realism is the only philosophy that doesn’t make the success of the science a miracle” (Putnam 1975a: 73).

One potential criticism of any indispensability argument is that by making the justification of our mathematical beliefs depend on our justification for believing in science, our mathematical beliefs become only as strong as our confidence in science. It is notoriously difficult to establish the truth of scientific theory. Some philosophers, like Nancy Cartwright and Bas van Fraassen have argued that science, or much of it, is false, in part due to idealizations. (See Cartwright 1983 and van Fraassen 1980.) The success of science may be explained by its usefulness, without presuming that scientific theories are true.

Still, even if science were only useful rather than true, PS claims that our mathematical beliefs may be justified by the uses of mathematics in science. The problems with scientific realism focus on the incompleteness and error of contemporary scientific theory. These problems need not infect our beliefs in the mathematics used in science. A tool may work fine, even on a broken machine. One could deny or remain agnostic towards the claims of science, and still attempt to justify our mathematical beliefs using Putnam’s indispensability argument.

The first two premises of PS are uncontroversial, so Putnam’s defense of PS focuses on its third premise. His argument for that premise is essentially a rejection of the argument that mathematics could be indispensable, yet not true. “It is silly to agree that a reason for believing that p warrants accepting p in all scientific circumstances, and then to add ‘but even so it is not *good enough*’” (Putnam 1971: 356).

For the Quinean holist, Putnam’s argument for PS3 has some force. Such a holist has no external perspective from which to evaluate the mathematics in scientific theory as merely useful. S/he can not say, “Well, I commit to mathematical objects within scientific theory, but I don’t really mean that they exist.”

In contrast, the opponent of PS may abandon the claim that our most sincere commitments are found in the quantifications of our single best theory. Instead, such an opponent might claim that only objects which have causal relations to ordinary physical objects exist. Such a critic is free to deny that mathematical objects exist, despite their utility in science, and nothing in PS prevents such a move, in the way that QI1-QI3 do for Quine’s original argument.

More importantly, any account of the applicability of mathematics to the natural world other than the indispensabilist’s refutes PS3. For example, Mark Balaguer’s plenitudinous platonism claims that mathematics provides a theoretical apparatus which applies to all possible states of the world. (See Balaguer 1998.) It explains the applicability of mathematics to the natural world, non-miraculously, since any possible state of the natural world will be described by some mathematical theory.

Similarly, and more influentially, Hartry Field has argued that the reason that mathematics is successful as the language of science is because it is conservative over nominalist versions of scientific theories. (See Field 1980, especially the preliminary remarks and Chapter 1.) In other words, Field claims that mathematics is just a convenient shorthand for a theory which includes no mathematical axioms.

In response, one could amend PS3 to improve Putnam's argument:

> PS3*: Realism best explains the success of mathematics as the language of science.

A defense of the new argument, PS*, would entail showing that realism is a better explanation of the utility of mathematics than other options.

§4: Resnik's Pragmatic Indispensability Argument

Michael Resnik, like Putnam, presents both a holistic indispensability argument, like Quine's, and a non-holistic argument called the pragmatic indispensability argument. In the pragmatic argument, Resnik first links mathematical and scientific justification.

RP
- RP1. In stating its laws and conducting its derivations, science assumes the existence of many mathematical objects and the truth of much mathematics.
- RP2. These assumptions are indispensable to the pursuit of science; moreover, many of the important conclusions drawn from and within science could not be drawn without taking mathematical claims to be true.
- RP3. So we are justified in drawing conclusions from and within science only if we are justified in taking the mathematics used in science to be true.
- RP4. We are justified in using science to explain and predict.
- RP5. The only way we know of using science thus involves drawing conclusions from and within it.
- RPC. So, by RP3, we are justified in taking mathematics to be true (Resnik 1997: 46-8).

RP, like PS, avoids the problems that may undermine our confidence in science. Even if our best scientific theories are false, their undeniable practical utility still justifies our using them. RP states that we need to presume the truth of mathematics even if science is merely useful. The key premises for RP, then, are the first two. If we can also take mathematics to be merely useful, then those premises are unjustified. The question for the proponent of RP, then, is how to determine whether science really presumes the existence of mathematical objects, and mathematical truth. How do we determine the commitments of scientific theory?

We could ask scientists about their beliefs, but they may work without considering the question of mathematical truth at all. Like PS, RP seems liable to the critic who claims that the same laws and derivations in science can be stated while taking mathematics to be merely useful. (See Azzouni 2004, Leng 2005, and Melia 2000.) The defender of RP needs a procedure for determining the commitments of science that blocks such a response, if not a more general procedure for determining our commitments. RP may thus be best interpreted as a reference back to Quine's holistic argument, which provided both.

§5. The Explanatory Indispensability Argument

Alan Baker and Mark Colyvan, in recent work, have defended an explanatory indispensability argument.

EI
- EI1. There are genuinely mathematical explanations of empirical phenomena.
- EI2. We ought to be committed to the theoretical posits postulated by such explanations.
- EIC. We ought to be committed to the entities postulated by the mathematics in question (Mancosu 2008: §3.2; see also Baker 2005 and Lyon and Colyvan 2008).

EI differs from Quine's original indispensability argument QI, and other versions of the argument, by seeking the justification for our mathematical beliefs in scientific explanations which rely on mathematics, rather than in scientific theories. Mathematical explanation and scientific explanation are difficult and controversial topics, beyond the scope of this article. Still, two comments are appropriate.

First, it is unclear whether EI is intended as a greater demand on the indispensabilist than the standard indispensability argument. Does the platonist have to show that mathematical objects are indispensable from both our best theories and our best explanations, or just from one of them? Conversely, must the nominalist dispense with mathematics from both theories and explanations?

Second, EI, like Putnam's success argument and Resnik's pragmatic argument, leaves open the question of how one is supposed to determine the commitments of an explanation. EI2 refers to the theoretical posits postulated by explanations, but does not tell us how we are supposed to figure out what an explanation posits. If the commitments of a scientific explanation are found in the best scientific theory used in that explanation, then EI is no improvement on QI. If, on the other hand, EI is supposed to be a new and independent argument, its proponents must present a new and independent criterion for determining the commitments of explanations.

Given the recent development of the explanatory indispensability argument and the current interest in mathematical explanation, it is likely that more work will be done on these questions soon.

§6. Characteristics of Indispensability Arguments in the Philosophy of Mathematics

Quine's indispensability argument relies on specific claims about how we determine our commitments, and what our canonical language must be. Putnam's success argument, Resnik's pragmatic argument, and the explanatory indispensability argument are more general, since they do not specify a particular method for determining the commitments of a theory. Putnam and Resnik maintain that we are committed to mathematics because of the ineliminable role that mathematical objects play in scientific theory. Proponents of the explanatory argument argue that our mathematical beliefs are justified by the role that mathematics plays in scientific explanations.

Indispensability arguments in the philosophy of mathematics can be quite general, and can rely on supposedly indispensable uses of mathematics in a wide variety of contexts. For instance, in later work, Putnam defends belief in mathematical objects for their indispensability in explaining our mathematical intuitions. (See Putnam 1994: 506.) Since he thinks that our mathematical intuitions derive exclusively from our sense experience, this later argument may still be classified as an indispensability argument.

Here are some characteristics of many indispensability arguments in the philosophy of mathematics, no matter how general:

> *Naturalism*: The job of the philosopher, as of the scientist, is exclusively to understand our sensible experience of the physical world.
>
> *Theory Construction*: In order to explain our sensible experience we construct a theory, or theories, of the physical world. We find our commitments exclusively in our best theory or theories.
>
> *Mathematization*: Some mathematical objects are ineliminable from our best theory or theories.
>
> *Subordination of Practice*: Mathematical practice depends for its legitimacy on natural scientific practice.

Although the indispensability argument is a recent development, earlier philosophers may have held versions of the argument. Mark Colyvan classifies arguments from both Frege and Gödel as indispensability arguments, on the strength of their commitments to Theory Construction and Mathematization. (See Colyvan 2001: 8-9.) Both Frege and Gödel, though, deny Naturalism and Subordination of Practice, so they are not indispensabilists according to the characterization in this section.

§7. Responses to the Indispensability Argument

The most influential approach to denying the indispensability argument is to reject the claim that mathematics is essential to science. The main strategy for this response is to introduce scientific or mathematical theories which entail all of the consequences of standard theories, but which do not refer to mathematical objects. Such nominalizing strategies break into two groups.

In the first group are theories which show how to eliminate quantification over mathematical objects within scientific theories. Hartry Field has shown how we can reformulate some physical theories to quantify over space-time regions rather than sets. (See Field 1980 and Field 1989.) According to Field, mathematics is useful because it is a convenient shorthand for more complicated statements about physical quantities. John Burgess has extended Field's work. (See Burgess 1984, Burgess 1991a, and Burgess and Rosen 1997.) Mark Balaguer has presented steps toward nominalizing quantum mechanics. (See Balaguer 1998.)

The second group of nominalizing strategies attempt to reformulate mathematical theories to avoid commitments to mathematical objects. Charles Chihara (Chihara 1990), Geoffrey Hellman (Hellman 1989), and Hilary Putnam (Putnam 1967b and Putnam 1975a) have all explored modal reformulations of mathematical theories. Modal reformulations replace claims about mathematical objects with claims about possibility.

Another line of criticism of the indispensability argument is that the argument is insufficient to generate a satisfying mathematical ontology. For example, no scientific theory requires any more than $\aleph_1$ sets; we don't need anything nearly as powerful as the full ZFC hierarchy. But, standard set theory entails the existence of much larger cardinalities. Quine calls such un-applied mathematics "mathematical recreation" (Quine 1986b: 400).

The indispensabilist can justify extending mathematical ontology a little bit beyond those objects explicitly required for science, for simplicity and rounding out. But few indispensabilists have shown interest in justifying beliefs in, say, inaccessible cardinals. (Though, see Colyvan 2007 for such an attempt.) Thus, the indispensabilist has a restricted ontology. Similarly, the indispensability argument may be criticized for making mathematical epistemology a posteriori, rather than a priori, and for classifying mathematical truths as contingent, rather than necessary. Indispensabilists may welcome these departures from traditional interpretations of mathematics. (For example, see Colyvan 2001, Chapter 6.)

8. Inter-theoretic and Intra-theoretic Indispensability Arguments

Indispensability arguments need not be restricted to the philosophy of mathematics. Considered more generally, an indispensability argument is an inference to the best explanation which transfers evidence for one set of claims to another. If the transfer crosses disciplinary lines, we can call the argument an inter-theoretic indispensability argument. If evidence is transferred within a theory, we can call the argument an intra-theoretic indispensability argument. The indispensability argument in the philosophy of mathematics transfers evidence from natural science to mathematics. Thus, this argument is an inter-theoretic indispensability argument.

One might apply inter-theoretic indispensability arguments in other areas. For example, one could argue that we should believe in gravitational fields (physics) because they are ineliminable from our explanations of why zebras do not go flying off into space (biology). We might think that biological laws reduce, in some sense, to physical laws, or we might think that they are independent of physics, or supervenient on physics. Still, our beliefs in some basic claims of physics seem indispensable to other sciences.

As an example of an intra-theoretic indispensability argument, consider the justification for our believing in the existence of atoms. Atomic theory makes accurate predictions which extend to the observable world. It has led to a deeper understanding of the world, as well as further successful research. Despite our lacking direct perception of atoms, they play an indispensable role in atomic theory. According to atomic theory, atoms exist. Thus, according to an intra-theoretic indispensability argument, we should believe that atoms exist.

As an example of an intra-theoretic indispensability argument within mathematics, consider Church’s Thesis. Church’s Thesis claims that our intuitive notion of an algorithm is equivalent to the technical notion of a recursive function. Church’s Thesis is not provable, in the ordinary sense. But, it might be defended by using an intra-theoretic indispensability argument: Church’s Thesis is fruitful, and, arguably, indispensable to our understanding of mathematics. Quine’s argument for QI2, that we must believe in the commitments of any theory we accept, might itself also be called an intra-theoretic indispensability argument.

§9. Conclusion

There are at least three ways of arguing for empirical justification of mathematics. The first is to argue, as John Stuart Mill did, that mathematical beliefs are about ordinary, physical objects to which we have sensory access. The second is to argue that while mathematical beliefs are about abstract mathematical objects, we have sensory access to such objects. (See Maddy 1990.) The currently most popular way to justify mathematics empirically is to argue:

A. Mathematical beliefs are about abstract objects;
B. We have experiences only with physical objects; and yet
C. Our experiences with physical objects justify our mathematical beliefs.

This is the indispensability argument in the philosophy of mathematics.

References

Azzouni, Jody. 2004. *Deflating Existential Consequence: A Case for Nominalism*. Oxford University Press.

Azzouni, Jody. 1998. "On 'On What There Is'." *Pacific Philosophical Quarterly* 79: 1-18.

Azzouni, Jody. 1997b. "Applied Mathematics, Existential Commitment, and the Quine-Putnam Indispensability Thesis." *Philosophia Mathematica (3)* 5: 193-209.

Baker, Alan. 2005. "Are there Genuine Mathematical Explanations of Physical Phenomena?" *Mind*: 114: 223-238.

Baker, Alan. 2003. "The Indispensability Argument and Multiple Foundations for Mathematics." *The Philosophical Quarterly* 53.210: 49-67.

Baker, Alan. 2001. "Mathematics, Indispensability and Scientific Practice." *Erkenntnis* 55: 85-116.

Balaguer, Mark. 1998. *Platonism and Anti-Platonism in Mathematics*. New York: Oxford University Press.

Balaguer, Mark. 1996. "Toward a Nominalization of Quantum Mechanics." *Mind* 105: 209-226.

Bangu, Sorin Ioan. 2008. "Inference to the Best Explanation and Mathematical Realism." *Synthese* 160: 13-20.

Benacerraf, Paul. 1973. "Mathematical Truth." In Paul Benacerraf and Hilary Putnam, eds., *Philosophy of Mathematics: Selected Readings*, second edition, Cambridge: Cambridge University Press, 1983.

Burgess, John. 1991b. "Synthetic Physics and Nominalist Realism." In C. Wade Savage and Philip Ehrlich, *Philosophical and Foundational Issues in Measurement Theory*, Hillsdale: Lawrence Erlblum Associates, 1992.

Burgess, John. 1991a. "Synthetic Mechanics Revisited." *Journal of Philosophical Logic* 20: 121-130.

Burgess, John. 1984. "Synthetic Mechanics." *Journal of Philosophical Logic* 13: 379-395.

Burgess, John. 1983. "Why I am Not a Nominalist." *Notre Dame Journal of Formal Logic* 24.1: 93-105.

Burgess, John, and Gideon Rosen. 1997. *A Subject with No Object*. New York: Oxford.

Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.

Chihara, Charles. 1990. *Constructabiliy and Mathematical Existence.* Oxford: Oxford University Press.

Colyvan, Mark. 2007. "Mathematical Recreation versus Mathematical Knowledge." In Leng, Mary, Alexander Paseau and Michael Potter, eds, *Mathematical Knowledge*, Oxford University Press, 109-122.

Colyvan, Mark. 2002. "Mathematics and Aesthetic Considerations in Science." *Mind* 11: 69-78.

Colyvan, Mark. 2001. *The Indispensability of Mathematics*. Oxford University Press.

Field, Hartry. 1989. *Realism, Mathematics, and Modality*. Oxford: Basil Blackwell.

Field, Hartry. 1980. *Science Without Numbers*. Princeton: Princeton University Press.

Frege, Gottlob. 1953. *The Foundations of Arithmetic*. Evanston: Northwestern University Press.

Gödel, Kurt. 1963. "What is Cantor's Continuum Problem?" In Paul Benacerraf and Hilary

Putnam, eds., *Philosophy of Mathematics: Selected Readings*, second edition. Cambridge: Cambridge University Press, 1983.

Gödel, Kurt. 1961. “The Modern Development of the Foundations of Mathematics in the Light of Philosophy.” In Solomon Feferman et al., eds., *Kurt Gödel: Collected Works, Vol. III*. New York: Oxford University Press, 1995.

Hellman, Geoffrey. 1989. *Mathematics Without Numbers*. New York: Oxford University Press.

Hintikka, Jaakko. 1973. “Quantifiers vs Quantification Theory.” *Dialectica* 27.3: 329-358.

Leng, Mary. 2005. “Mathematical Explanation.” In Cellucci, Carlo and Donald Gillies eds, *Mathematical Reasoning and Heuristics*, King's College Publications, London, 167-189.

Lyon, Aidan and Mark Colyvan. 2007. “The Explanatory Power of Phase Spaces.” *Philosophia Mathematica* 16.2: 227-243.

Maddy, Penelope. 1992. “Indispensability and Practice.” *The Journal of Philosophy* 89: 275-289.

Maddy, Penelope. 1990. *Realism in Mathematics*. Oxford: Clarendon Press.

Mancosu, Paolo. “Explanation in Mathematics.” *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/mathematics-explanation/>.

Marcus, Russell. 2007. “Structuralism, Indispensability, and the Access Problem.” *Facta Philosophica* 9, 2007: 203-211.

Melia, Joseph. 2002. “Response to Colyvan.” *Mind* 111: 75-79.

Melia, Joseph. 2000. “Weaseling Away the Indispensability Argument.” *Mind* 109: 455-479.

Melia, Joseph. 1998. “Field’s Programme: Some Interference.” *Analysis* 58.2: 63-71.

Mendelson, Elliott. 1997. *Introduction to Mathematical Logic*, 4th ed. Chapman & Hall/CRC.

Mill, John Stuart. 1941. *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. London, Longmans, Green.

Putnam, Hilary. 1994. “Philosophy of Mathematics: Why Nothing Works.” In his *Words and Life*. Cambridge: Harvard University Press.

Putnam, Hilary. 1975b. *Mathematics, Matter, and Method: Philosophical Papers, Vol. I.* Cambridge: Cambridge University Press.

Putnam, Hilary. 1975a. “What is Mathematical Truth?” In Putnam 1975b.

Putnam, Hilary. 1974. “Science as Approximation to Truth.” In Putnam 1975b.

Putnam, Hilary. 1971. *Philosophy of Logic*. In Putnam 1975b.

Putnam, Hilary. 1967b. “Mathematics Without Foundations.” In Putnam 1975b.

Putnam, Hilary. 1967a. “The Thesis that Mathematics is Logic.” In Putnam 1975b.

Putnam, Hilary. 1956. “Mathematics and the Existence of Abstract Entities.” *Philosophical Studies* 7: 81-88.

Quine, W.V. 1995. *From Stimulus to Science*. Cambridge: Harvard University Press.

Quine, W.V. 1986b. “Reply to Charles Parsons.” In Lewis Edwin Hahn and Paul Arthur Schilpp, eds., *The Philosophy of W.V. Quine*. La Salle: Open Court, 1986.

Quine, W.V. 1986a. *Philosophy of Logic*, 2nd edition. Cambridge: Harvard University Press.

Quine, W.V. 1981. *Theories and Things*. Cambridge: Harvard University Press.

Quine, W.V. 1980. *From a Logical Point of View*. Cambridge: Harvard University Press.

Quine, W.V. 1978. “Success and the Limits of Mathematization.” In Quine 1981.

Quine, W.V. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University

Press.
Quine, W.V. 1960. *Word & Object*. Cambridge: The MIT Press.
Quine, W.V. 1958. "Speaking of Objects." In Quine 1969.
Quine, W.V. 1955. "Posits and Reality." In *The Ways of Paradox*. Cambridge: Harvard University Press, 1976.
Quine, W.V. 1951. "Two Dogmas of Empiricism." In Quine 1980.
Quine, W.V. 1948. "On What There Is." In Quine 1980.
Quine, W.V. 1939. "Designation and Existence." In Feigl and Sellars, *Readings in Philosophical Analysis*, Appleton-Century-Crofts, Inc., New York: 1940.
Resnik, Michael. 1997. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press.
Resnik, Michael. 1995. "Scientific vs. Mathematical Realism: The Indispensability Argument. *Philosophia Mathematica (3)* 3: 166-174.
Resnik, Michael D. 1993. "A Naturalized Epistemology for a Platonist Mathematical Ontology." In Sal Restivo, et. al., eds., *Math Worlds: Philosophical and Social Studies of Mathematics and Mathematics Education*, Albany: SUNY Press, 1993.
Shapiro, Stewart. *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford University Press, 1991.
Sher, Gila. 1990. "Ways of Branching Quantifiers." *Linguistics and Philosophy* 13: 393-422.
Sober, Elliott. 1993. "Mathematics and Indispensability." *The Philosophical Review* 102: 35-57.
Tarski, Alfred. 1944. "The Semantic Conception of Truth: and the Foundations of Semantics." *Philosophy and Phenomenological Research* 4.3: 341-376.
Van Fraassen, Bas C. 1980. *The Scientific Image*. Oxford: Clarendon Press.
Whitehead, Alfred North and Bertrand Russell. 1997. *Principia Mathematica to *56*. Cambridge University Press.

Some Entries for *Key Terms in Logic*
Russell Marcus
Hamilton College

Published in *Key Terms in Logic*, Jon Williamson and Federica Russo, Continuum Publishers, November 2010.

Model Theory (491):

All communication involves the production and interpretation of statements. Semantics studies the interpretations of statements. The study of the statements of a formal system is called metatheory. Metatheory may be divided into proof theory and model theory. Proof theory studies the rules guiding inferences within the system. Model theory is a mathematical approach to semantics, in particular to the assignment of truth values to the statements of a theory.

A formal system consists of a language (vocabulary and formation rules) as well as axioms and rules for generating theorems. Given a formal system, the first step in model theory is to specify an interpretation of each particle of the system. Then, we provide rules governing the assignments of truth values to complex expressions on the basis of assignments of truth values to their component parts. A model is an interpretation of a system on which its theorems are true.

The semantics for the propositional calculus are easily given without model theory. Truth tables suffice to interpret the connectives, and propositional variables can be replaced by propositions or sentences.

The semantics for predicate logic normally proceeds using set theory. We specify a domain of interpretation for the variables of the system. For example, it is natural to use the domain of natural numbers to model the Peano Axioms, and to use sets to model the axioms of set theory. Models of physical theories naturally take the physical world as their domains. Non-standard models, using unintended domains of quantification, are available.

The next step in constructing a model is to assign elements of the domain to particles of the system. We assign particular objects to the constants. Predicates are normally interpreted as sets of objects in the domain, and n-place relations are taken as ordered n-tuples within the domain. An existentially quantified sentence is true in a model if there is an object in the domain of interpretation with the properties mentioned in the sentence. A universally quantified expression is true if the properties mentioned hold of every object in the domain.

For modal logics, Kripke models provide possible-worlds semantics. In a Kripke model, we start with a set of ordinary models, one for each possible world, and an accessibility relation among them. A statement is taken to be possible if there is an accessible possible world in which the statement is true. A statement is taken to be necessary if it is true in all possible worlds.

Model theory, developed in large part by Alfred Tarski and Abraham Robinson in the mid-twentieth century, has become a standard tool for studying set theory and algebraic structures. Major results of model theory include Paul Cohen's proof that the continuum hypothesis is independent of the axioms of Zermelo-Fraenkel (ZF) set theory including the axiom of choice, and that the axiom of choice itself is independent of the other axioms of ZF. Model theory is responsible for the so-called Skolem paradox, one of its earliest results.

See also: completeness, metalanguage, soundness, Hilbert

Semantic Tree (284):

In propositional logic, we can determine whether a set of formulas is consistent by examining the truth tables for the set. Alternatively, we can construct a semantic tree. Semantic trees may be used to test an argument for validity, since an argument is invalid if, and only if, the negation of its conclusion is consistent with the truth of its premises. Semantic trees are less cumbersome than truth tables, providing an easy method for testing a large set of formulas. They require less creative construction than natural deductions. There are decision procedures for semantic trees for propositional logic, which means that the procedure will always terminate in a solution.

To construct a semantic tree, we replace compound formulas with simpler sub-formulas with the same truth conditions. For example, we can replace ‘$\neg(A \vee B)$’ with ‘$\neg A$’ and ‘$\neg B$’, since the longer formula is true if, and only if, the shorter ones are true. Some replacement rules branch, giving the construction the appearance of a tree. For example, any branch on which ‘$A \rightarrow \neg B$’ appears divides into two branches, one which contains ‘$\neg A$’ and the other which contains ‘B’. The tree is completely constructed when all compound formulas have been replaced by either simple formulas or negations of simple formulas. If all branches contain a contradiction (a simple propositional formula and its negation) then the original set of formulas is consistent.

Semantic trees are useful in predicate logic, as well. For sets of formulas with only monadic predicates, the rules determine a decision procedure. There is no decision procedure for some sets of formulas with relational predicates. Semantic trees are useful in modal logic.

Semantic trees are also called truth trees, semantic tableau, or semantic tableaux.

See also: deduction, proof

Syntax (50):

Providing the syntax of a formal theory is the first step in its construction. The syntax consists of the language, or alphabet, in which the theory is written, along with rules for constructing well-formed formulas. Syntax is opposed to semantics, which governs the assignment of truth values to those formulas.

Variable (48):

Variables in formal systems function similarly to pronouns in natural language; their reference varies. Variables stand for other logical particles. Sentences or propositions may substitute for propositional variables. Objects or names of objects substitute for variables in first-order predicate logic. Higher-order logics also contain variables in predicate positions.

Church (391):

In 1936, Alonzo Church, Emil Post, and Alan Turing each proposed independent explications of the informal notion of an effectively computable function, or algorithm. The three formal notions were later shown to select the same class of mathematical functions. Further equivalent formulations have been produced by Gödel and others. The resulting thesis, that the computable functions are the recursive functions, has become known as Church's Thesis, or the Church-Turing thesis. (The notion of a recursive function traces to Gödel; Church considered a related class of functions called λ-definable.)

Church's Thesis is important because we want to know whether some problems have algorithmic solutions. For example, Church initially formulated the thesis in an attempt to answer the question of whether first-order logic was decidable. A theory is decidable if there is a procedure for determining whether any given formulais a theorem. Since recursion is formally definable, Church's Thesis provides a method for determining whether a particular problem has an effective solution. It provides a formal characterization of an intuitive concept.

Church's Thesis is also important because of its relation to Turing's formulation. Turing selected the recursive functions by considering the abilities of logical computing machines, or Turing Machines. Some writers who have compared Turing machines to human minds have used Church's Thesis with excessive enthusiasm, making broader claims about its implications than are supported by the thesis properly construed. In contrast, Church's Thesis is entirely silent about the nature and limitations of both the human mind and computing machines.

Church's Thesis is widely accepted. It appears that every recursive function is effectively computable, and it also appears that every effectively computable function is recursive. Still, there is some debate over whether Church's Thesis is provable. This debate has focused on whether any identification of an informal concept with a formal notion can be proven. Some philosophers consider Church's Thesis to be a working hypothesis. Others take it to be merely another mathematical refinement of a commonsense notion like set, function, limit, or logical consequence.

Church's Thesis is independent of the purely technical result called Church's Theorem. Church's Theorem shows that first-order logic is recursively undecidable, as long as it contains non-monadic predicates. (A monadic predicate, like 'is blue' takes only one variable, in contrast to relational predicates, like 'is bigger than' or 'is between'.) Appending Church's Thesis to Church's Theorem, as Church did, we can show that there is no effective decision procedure for first-order logic, no sure method for deciding whether a formula of first-order logic is valid. Though, there is a decision procedure for monadic predicate logic.

See also: computability, decidability

Hilbert (496):

David Hilbert, 1862-1943, should be viewed as the progenitor of metatheory in logic and mathematics. He was among the most prominent mathematicians of his time. His achievements in mathematics include work on geometry, number theory, algebra and analysis; he also contributed to the theory of relativity. Hilbert shaped the direction of mathematical thought in the twentieth century, most famously by framing the Paris Problems: twenty-three open questions presented at the 1900 International Congress of Mathematicians.

In philosophy of mathematics, Hilbert is variously characterized as a formalist and as a finitist. While Hilbert's views contain elements of both formalism and finitism, neither of these terms effectively captures the subtlety of his thought. Some formalists hold that mathematical theories are best understood as uninterpreted systems. Some finitists reject all infinitary results. In contrast, Hilbert believed both that some mathematical statements were true of real objects and that transfinite mathematics was legitimate.

In the early twentieth century, philosophers of mathematics struggled to understand the ramifications of various oddities of set theory, including Cantor's paradox (arising from consideration of the set of all sets), the Burali-Forti paradox (arising from consideration of the well-ordering of the ordinals), and Russell's paradox (arising from the assumption that every property determines a set). Intuitionists, e.g. Brouwer, concluded that the infinitary mathematics which leads to these paradoxes was illegitimate. Hilbert, in contrast, wished to establish finitistic foundations for infinitary mathematics.

Hilbert distinguished between real and ideal mathematical formulas. Real formulas are generally finitistic, and may be directly verified. Mathematical theories which included ideal elements were instead to be tested for their consistency. Unlike logicists like Frege, for whom the consistency of mathematics follows from the presumed truth of its axioms, Hilbert took the consistency of a set of axioms as sufficient evidence for mathematical legitimacy. Further, Hilbert took ideal formulas to be meaningless. Hilbert's emphasis on consistency and his claims about the meanings of ideal formulas have led people to consider him a formalist.

In addition to consistency, if one can show completeness, that every valid theorem is provable, then one could hope for a solution to all open mathematical problems. Hilbert tried to establish that mathematical theories were both consistent and complete by studying mathematical systems themselves. He thus founded the metamathematics and metalogic that characterize much of contemporary logical research.

Many logical theories are provably consistent and complete. In contrast, Gödel's incompleteness theorems struck a decisive blow against Hilbert's pursuits these results for mathematics. Gödel's first theorem showed that, for even quite weak mathematical theories, a consistent theory could not be complete. Gödel's second theorem proved that the consistency of a theory could never be proven within the theory itself. We can only prove that mathematical theories are consistent relative to other theories.

Hilbert's views survive in Hartry Field's fictionalism, which emphasizes the consistency of mathematical theories; in Mark Balaguer's plenitudinous platonism, which asserts that every consistent mathematical theory truly describes a mathematical realm; and in defenses of limited versions of Hilbert's Programme.

See also: intuitionism, logicism

Leopold Löwenheim (49):

Leopold Löwenheim, 1878-1957, is known for his work on relations and for an early negative result in logic, which has come to be known as the Löwenheim-Skolem theorem. The theorem says that if a first-order theory requires models of infinite size, models can be constructed of any infinite size.

See also: model

Thoraf Skolem (50):

Thoraf Skolem, 1887-1963, simplified Löwenheim's theorem which states that if a first-order theory requires models of infinite size, models can be constructed of any infinite size. 'Skolem's paradox' refers to the observation that first-order theories which yield theorems asserting the existence of uncountably many (or more) objects have denumerable models.

See also: model

The Explanatory Indispensability Argument

Work in Progress

Russell Marcus

Department of Philosophy, Hamilton College

198 College Hill Road

Clinton NY 13323

rmarcus1@hamilton.edu

(315) 859-4056 (office)

(315) 381-3125 (home)

Abstract:

A new version of the indispensability argument in the philosophy of mathematics may be found, especially in work by Mark Colyvan and Alan Baker. The new argument relies in part on recent research in mathematical explanation, and alleges that our mathematical beliefs are justified by their indispensable appearances in scientific explanations. This paper provides an analysis of the new argument, as well as a brief survey of some earlier versions of the indispensability argument which share some of the new argument's features. Like some earlier versions, the explanatory indispensability argument is unfortunately elliptical about a criterion for ontological commitment, and so liable to an instrumentalist, or fictionalist, response.

There are two overlapping arguments in this paper. The first concludes that explanations which increase our understanding may not be those in which we express our ontological commitments. The second concludes that what has become known as easy-road, or weasel, nominalism is an appropriate response only to some indispensability arguments, and not to Quine's original argument, which was designed to block such weaseling. The explanatory indispensability argument is no improvement on the Quinean one.

§1: Explanations and Theories

Q: Should the following inference convince us that there are numbers?

IM I have two mangoes.
Andrés has three different mangoes.
So, together we have five mangoes.

A: No, because simple, adjectival uses of arithmetic are easily eliminated.

IN $(\exists x)(\exists y)(Mx \bullet My \bullet Bxm \bullet Bym \bullet x \neq y)$
$(\exists x)(\exists y)(\exists z)(Mx \bullet My \bullet Mz \bullet Bxa \bullet Bya \bullet Bza \bullet x \neq y \bullet x \neq z \bullet y \neq z)$
$(x)[(Mx \bullet Bxa) \supset {\sim}Bxm]$
$\therefore (\exists x)(\exists y)(\exists z)(\exists w)(\exists v)(Mx \bullet My \bullet Mz \bullet Mw \bullet Mv \bullet x \neq y \bullet x \neq z \bullet x \neq w \bullet x \neq v \bullet y \neq z \bullet y \neq w \bullet y \neq v \bullet z \neq w \bullet z \neq v \bullet w \neq v)$[1]

The contrast between IM, the mathematical inference, and IN, the parallel nominalistic inference, demonstrates that some statements which use numbers may be taken as convenient shorthand for more-complicated statements that do not. There may be statements more complex than those in IM from which mathematical objects are ineliminable. Indeed, the last thirty years of philosophy of mathematics has been centrally concerned, in one way or another, with this question. But the eliminability of numbers is uncontroversial in some cases. If someone were to present IM as a reason for believing that there are numbers, anyone would be justified in denying the inference to the existence of mathematical objects by proffering IN as a replacement. Whether we believe in mathematical objects or not, IM is not a good reason for believing in them.

IM can not justify beliefs in mathematical objects because it is a casual inference which does not reflect our serious commitments. When we want to display our ontological commitments, we speak most soberly, invoking parsimony and rewriting our casual sentences. We reflectively remove from our language references to sakes and behalves and point-masses and frictionless planes, and to adjectival uses

[1] A more heavy-handed option to remove numbers from IM would be to replace the terms which refer to them with references to ordered regions of substantivalist space-time; see Field 1980 and Burgess and Rosen 1997: §IIA.

of natural numbers.[2] IN makes it clear that the essential subjects of IM are mangoes, not numbers.

Now, consider the question, "Why are there five mangoes here?" A good explanation of the fact that there are five mangoes on the table is that I brought two and Andrés brought three. That fact is explained by IM, and only awkwardly demonstrated, if explained at all, by IN. IM is not a complete, best theory of mangoes. It requires background assumptions about object constancy, and that mangoes do not annihilate each other when, say, there are more than three together. But, it is an explanation that will satisfy any ordinary person, and even a philosopher who is not thinking too hard about theories of explanation. The only way for IN to have any explanatory force for someone is to translate it back to something like IM.

Both IM and IN have their virtues. IM provides a perspicuous and easily-understood explanation. The conclusion of IN follows from its premises in first-order logic by purely computational means. A philosopher may insist that the first-order derivability of its conclusion makes IN more explanatory. Indeed, IN is the kind of inference on which the traditional Deductive-Nomological (D-N) model of explanation relies.

On the D-N, or covering-law, account, the explanation of a state of affairs is a logical inference, involving the laws of a serious theory combined with appropriate initial conditions. The theories to which D-N explanations appeal are ones in which we speak most strictly. Modifications of the D-N model, like Railton's model of probabilistic explanation, or Kitcher's unificationist account, work similarly. Kitcher, for example, invokes unifying argument patterns which also answer why-questions by proffering inferences made within a serious theory. While IN is a simple logical inference which suppresses auxiliary presuppositions involving laws governing mangoes, we could easily tidy it up, appealing to the relevant laws of conservation of mangoes, and such. IN could thus play a central role in a D-N, or related, explanation of why there are five mangoes here.

[2] On sakes and behalves, see Quine 1960: 244.

So, there are two distinct senses of 'explanation' on the table. The first sense is represented by the casual inference IM. In contrast, the D-N and related models of explanation capture the relation between laws and explanations, perhaps the most salient feature in common to a wide range of explanations. For reasons that will become clear, I focus in this paper on the first sense. There is a new version of the indispensability argument in the philosophy of mathematics promoted as an alternative to the traditional Quinean argument. The new explanatory indispensability argument puts aside the question of whether mathematical objects are eliminable from our best, most serious theories, and focuses on whether mathematics plays an indispensable role in explanations. If the explanatory argument were to rely on a D-N (or related) model of explanation, it would not differ from the Quinean argument. The central question about the argument would be whether mathematical objects could be eliminated from our best theories, as IN eliminates the references to numbers in IM. Thus, the proponents of the explanatory argument must appeal to a different model of explanation.

Let's call the kind of explanation that IM provides, but that IN does not provide, epistemic, in contrast to the D-N and related kinds of explanation, which we can call metaphysical. An epistemic explanation is used to increase understanding.[3] When we invoke metaphors in science, or use models for explanation, we rely on the epistemic notion. We do not think that the atom is literally constructed like the solar system; nevertheless, the image of electron orbits can be a useful heuristic, as long as it is not taken too seriously.

Metaphysical explanations involve covering laws or unifying principles and initial conditions. They rely on arguments in which we speak most soberly, like IN. To account for residual concerns that

[3] Pragmatic accounts of explanation have epistemic elements, as do causal-relevance accounts. See Friedman 1974 on the failure of D-N accounts to emphasize understanding. Some metaphysical accounts, like Kitcher's, do attempt to assimilate inference and understanding. Still, the influence of Hempel's covering-law model, and its attempt to provide an objective account of explanation extends to more-recent work. Railton, for example, promises, "An account of probabilistic explanation *free from relativization to our present epistemic situation*" (Railton 1978: 219, emphasis added).

explanations should foster understanding, the proponent of a metaphysical account may promise that when we understand the laws or causal structures or unifying principles underlying a phenomenon, we will understand why it occurred.

Both senses of 'explanation' have legitimate uses. This paper takes no position on whether our ordinary notion of explanation is sufficiently captured by any metaphysical or epistemic account. The points I want to make require only that we can distinguish between a sense of 'explanation' which depends exclusively on our most parsimonious theories and another sense, which underlies the new indispensability argument, for which our best theories do not suffice. I have illustrated the two kinds of explanations: IM is not metaphysically explanatory, but is epistemically explanatory. IN is metaphysically explanatory, but not epistemically so.

There are two morals to be learned from the differences between IM and IN.

Moral 1: We are committed to mathematical objects not by our casual uses of numbers, but only when we are speaking most seriously.
Moral 2: The theory we use to specify our ontological commitments may not be most useful when we want to explain facts about the world, in the epistemic sense of 'explain'.

The central claim of this paper is that the explanatory indispensability argument is no improvement on the Quinean argument because it depends on an equivocation between the two senses of 'explanation'. If it appeals to an epistemic notion of explanation, then our uses of mathematics should not be taken seriously. If it appeals to a metaphysical notion, then it is just a restatement of the original argument.

There are a variety of versions of the indispensability argument available, and a slew of different responses. The criticism I present in this paper applies to several different arguments, so it will be helpful to examine some of them. I start with a weak version.

§2: The Mathematical-Practice Argument

Any indispensability argument which purports to justify mathematical beliefs must specify when we are speaking most seriously. The chips were not down in IM, but that was an easy case. The more general problem of knowing precisely when we are speaking most seriously and revealing our true ontological commitments is trickier.

Consider this recent exchange: Jodi Azzouni claims that we can distinguish among our commitments to a theory's posits. We have observational access to thick posits and excuses for not being able to observe thin posits like objects outside of our light cone. So, we should believe in the thick and thin posits of a theory. In contrast, Azzouni claims that we should not believe in the existence of posits which we can call very thin, the result of mere causal talk (Azzouni 2004: Chapter 6). Mark Colyvan responds: "Whether [mathematical objects] are thin or very thin depends on what can count as an excuse for not being accessed thickly... Unfortunately Azzouni doesn't give us any guidance; he offers no systematic story about acceptable excuse clauses" (Colyvan 2010: 7). Colyvan is asking for a general account of when we are speaking seriously.

To take another particular case, we might wonder whether we are speaking most seriously in pure mathematics. The indispensability argument is supposed to be consistent with naturalist, even empiricist, principles.[4] One way of naturalizing mathematical epistemology would be to take the claims of practicing mathematicians at face value. Mathematicians say things like, "There are infinitely many primes." If we take such claims at face value, mathematical objects exist. Call this argument the mathematical-practice argument.

Alan Baker, calling deference to the practice of mathematicians strong mathematical naturalism, observes that such deference renders the indispensability argument moot (Baker 2003: 63-4). The

[4] See, for example, the mere title of Resnik's "A Naturalized Epistemology for a Platonist Mathematical Ontology" (Resnik 1983).

proponent of the traditional indispensability argument wonders whether mathematical objects are required for science. But, mathematical objects are required for face-value interpretations of mathematics. Thus, the defender of the mathematical-practice argument has no need to wonder about our scientific theories or explanations or practices.

The central problem with the mathematical-practice argument is the same as the problem with reading commitments from IM. We lack justification for taking mathematical practice itself seriously. Even if we accept mathematical practice generally, the references to mathematical objects by mathematicians in their work may be unserious. The work of Geoffrey Hellman (Hellman 1989) and Charles Chihara (Chihara 1990) in reformulating mathematical claims as modal claims is meaningful precisely because we need not take mathematical claims at face value. Michael Potter extends the point: "What mathematicians *say* is not always a reliable guide to what they are doing: what they mean and what they say they mean are not always the same" (Potter 2007: 18).

Baker contrasts strong naturalism with Penelope Maddy's weak mathematical naturalism. According to Maddy, ontological questions are external to the practice of mathematics and the indispensability argument, focusing on serious claims in science, is in play. Of course, scientists may be no better than mathematicians in saying precisely what they mean. As we will see in the next section, Joseph Melia argues that we can interpret both mathematicians and scientists as taking back all *prima facie* commitments to abstracta. "It is quite common for both scientists and mathematicians to think that their everyday, working theories are only partially true" (Melia 2000: 457).

If we want to know whether the use of some theory or discourse commits us to the existence of mathematical objects, and we want to avoid begging the question, we need an independently-motivated account of when we are speaking most seriously.

§3: Enter the Weasel

The strongest indispensability argument, at which the nominalist strategies of Field 1980 and others are aimed, is Quine's, because he takes seriously the problem of knowing when the chips are down. Though Quine's argument is well-known, I want to display it carefully. Key details of the argument are sometimes overlooked and this neglect has opened the argument to inappropriate criticism. In this section, I discuss how one interpretation of the indispensability argument weakens it. Then, I will present Quine's argument with its full armor.

Quine himself is at least partly at fault for some of the confusion surrounding the indispensability argument, since he never presented a specific, detailed version of the argument. Some philosophers see its reliance on holism as its central characteristic.[5] But, any indispensability argument has to give holism some role. The indispensabilist's claims that evidence for our scientific theory extends to the mathematical elements of that theory. Without some measure of holism, it is difficult to make the case that evidence transfers from science to mathematics.

The feature of Quine's argument which distinguishes it from other versions is Quine's insistence on specifying the details of how and when we are to be taken as speaking most seriously about our ontological commitments. These details arise out of a combination of his holism and his naturalism, as well as his criterion for determining our ontological commitments. The broad way in which the indispensability argument is generally represented masks these central claims. For example, Colyvan's version suppresses Quine's criterion for ontological commitment:

---

[5] See Baker 2005: 224 and Melia 2000: 455-6. By 'holism' in this paper, I refer to what is ordinarily known as confirmation holism, and which is the only holism directly relevant to the indispensability argument. Quine famously argues for confirmation holism from the stronger and more controversial semantic holism, especially in "Two Dogmas of Empiricism" (Quine 1951). But there are simpler argument for confirmation holism. See Colyvan 2001: §2.5.

1. We ought to have ontological commitment to all and only those entities that are indispensable to our best scientific theories.
2. Mathematical entities are indispensable to our best scientific theories.

Therefore:

3. We ought to have ontological commitment to mathematical entities (Colyvan 2001: 11).

Colyvan's first premise does not answer the question of when an entity is indispensable to our best theory: How does a theory make its posits? How do we read those posits? Are all the posits made in the same way, with the same force?[6]

A nominalist can respond to Colyvan's version of the argument by accepting, say, that vectors in Hilbert space are indispensable to the practice of quantum mechanics, but adding that we can, when speaking most seriously and parsimoniously, deny that our best theory really posits them. Indeed, Melia claims that one can "weasel out" of the indispensability argument, accepting that mathematics is ineliminable from scientific theory but maintaining that we need not believe that mathematical objects exist.[7] Melia defends weaseling by claiming that scientists use mathematics in order to express facts that are not representable without mathematics. But, such representations are not supposed to be ontologically serious. "The mathematics is the necessary scaffolding upon which the bridge must be built. But once the bridge has been built, the scaffolding can be removed" (Melia 2000: 469).[8]

Melia provides a helpful analogy for those resistant to the weasel response. Consider the two-

---

[6] Similarly, see Balaguer 1998: §2.1.

[7] See Melia 2000. Other recent versions of the weaseling strategy include Balaguer 1998, Chapter 7; Yablo 2005; Leng 2002; Leng 2005; Pincock 2004a and Pincock 2004b. See Colyvan 2010 and Daly and Langford 2010 for recent discussions. Rudolph Carnap was a proto-weasel; see Carnap 1950: 215.

[8] Mark Balaguer, Mary Leng, and Chris Pincock make similar remarks concerning the ontologically casual role of mathematics as a tool for expressing, representing, and modeling. For example, "We are not committed to belief in the existence of objects posited by our scientific theories *if their role in those theories is merely to represent configurations of physical objects*. Fictional objects can represent just as well as real objects can" (Leng 2005: 179). See also Balaguer 1998 and Pincock 2004b.

dimensional surface of a sphere. From a three-dimensional perspective, we can describe the surface as the locus of points equidistant from the center. But, the center is not part of the two-dimensional, non-Euclidean surface. From the point of view of the surface of the sphere, we can both appeal to the center while not really taking it as part of the world. "We do successfully and unproblematically describe a particular non-Euclidean world by taking back some of the implications of what we earlier said" (Melia 2000: 468).[9]

Note that there are two entirely different kinds of negative responses to the indispensability argument. The dispensabilist strategy is to eliminate our beliefs in mathematical objects by rewriting our claims which refer to mathematical objects as ones which do not refer to mathematical objects. The replacement of IM by IN illustrates a simple dispensabilist strategy. Field's reformulation of Newton's theory of gravity is a more sophisticated version. The second negative response to the indispensability argument, weaseling, just issues a flat denial of any commitments to mathematical objects, whether or not dispensabilist alternatives to our theories are available. Colyvan 2010 calls the dispensabilist strategy hard-road nominalism and the weasel strategy easy-road nominalism.

It is not my goal in this paper to defend or criticize the weaseling strategy. My claim is merely that certain versions of the indispensability argument are particularly liable to weaseling, while others resist the weasel. I will note, though, that a proponent of weaseling who provides an alternative criterion of ontological commitment, like Azzouni, who defends a version of the eleatic principle, is in a better position than one, like Melia, who merely denies commitments to mathematical objects.

[9] Daly and Langford 2010: 1114 point out that the surface can be described without appealing to the center of the sphere. Such a description will be less attractive, though.

§4: Quine's Indispensability Argument

Quine's version of the indispensability argument resists weaseling by stating precisely how to determine the commitments of a theory. Quine's argument not only demands that we transfer evidence from science to mathematics, and that our scientific theories are the locus of our real commitments, but also that we find our commitments in a particular way.

QI

QI1. We should believe the theory which best accounts for our sense experience.
QI2. If we believe a theory, we must believe in its ontological commitments.
QI3. The ontological commitments of any theory are the objects over which that theory first-order quantifies.
QI4. The theory which best accounts for our sense experience first-order quantifies over mathematical objects.
QIC. We should believe that mathematical objects exist.[10]

QI tells us that the chips are down in our best scientific theory. It does not tell us how to determine which theory is our best. But according to Quine's naturalism, criteria for ranking theories, the super-empirical virtues, are matters for scientists to determine. More importantly, QI tells us precisely how to determine the commitments of our best theory. Premises QI2 and QI3 are vital because they block attempts to weasel out of the indispensability argument. When the weasel says that we can differentiate between the real and the merely instrumental posits of our theory, Quine's holism blocks the move: all posits are on par. When the weasel says that we can take back what we allege, Quine's naturalism denies that such double-talk is sensible or defensible. QI derives its strength from Quine's connection between ontology and the construction of formal theories. "The quest of a simplest, clearest overall pattern of canonical notation is not to be distinguished from a quest of ultimate categories, a limning of the most general traits of reality" (Quine 1960: 161).

The double-talk criticism is essential to QI. For Quine, if our best theory requires electrons for its bound variables, then we should believe in electrons. If it requires sets, we are committed to sets.

---

[10] See Quines 1939, 1948, 1951, 1955, 1958, 1960, 1978, and 1986.

Quine's response to Carnap's internal/external distinction, for example, relies on the double-talk criticism. Once one has accepted mathematical objects as an internal matter, one can not merely dismiss these commitments as the arbitrary, conventional adoption of mathematical language. Similarly, Quine's response to the Meinongian Wyman in "On What There Is," who presents two species of existence, is a variation on the double-talk criticism. We must distinguish between the meaningfulness of 'Pegasus' and its reference in order to avoid admitting that Pegasus subsists while at the same time denying that Pegasus exists. Putnam, defending Quine's argument, makes the double-talk criticism explicitly. "It is silly to agree that a reason for believing that p warrants accepting p in all scientific circumstances, and then to add 'but even so it is not *good enough*'" (Putnam 1971: 356).

Worries about double-talk bother Quine's critics as well as his friends. Field applies the double-talk criticism directly to mathematics. "If one *just* advocates fictionalism about a portion of mathematics, without showing how that part of mathematics is dispensable in applications, then one is engaging in intellectual doublethink..." (Field 1980: 2).

Proponents of the weaseling strategy depend precisely on the legitimacy of double-talk, of taking back what one initially alleges. "In order to communicate his picture of the world, [the weasel-nominalist] Joe must make clear that, as far as he is concerned, [the platonistic scientific theory] T* is wrong in certain respects. What Joe wants to do is *subtract* or *prune away* the platonistic entities whose existence is entailed by T*. Accordingly, in communicating his belief about the world, Joe might say, 'T* - but there are no such things as sets. In doing so, he is asserting some sentences, whilst denying one of their logical consequences" (Melia 2000: 466-7).

An indispensability argument which rejects the legitimacy of double-talk is resistant to weaseling. QI says that the commitments of a theory are to be found exclusively and uniformly in its quantifications. Once we differentiate among the posits of a theory, we are rejecting the cornerstones of QI which block the weasel: the combination of Quine's holism, his naturalism, and his criterion for ontological commitment.

QI is not resistant to all criticism. One sort of criticism attacks QI4. Field 1980 attempted to rewrite Newtonian Gravitational Theory (NGT) by quantifying over space-time regions rather than real numbers. Burgess's later refinements gave hope to nominalists. This hope was not unbridled, and the projects received significant criticism. The current consensus about dispensabilist projects is that something pretty close to Field's project can work for NGT, but that other theories, including those based on curved space-time and those which rely on statistical frameworks, are resistant. Still, advances have been made, like Balaguer's steps toward nominalizing quantum mechanics, and, as Burgess and Rosen argue, the lack of dispensabilist strategies currently available is weak evidence for their eventual non-existence.[11] No neat, first-order theory which eschews mathematical axioms will suffice for all of current and future science. But the dispensabilist has reasonable hope of finding moderately-attractive reformulations of large swaths of scientific theory.

Another sort of criticism of QI attacks Quine's holism.[12] Though such criticisms are beyond the range of this paper, it is important to note that, along with dispensabilist constructions, they are responsible for the new, explanatory indispensability argument. "It would be useful to have a version of the Indispensability Argument that did not rely so crucially on holism" (Baker 2005: 224) Before we turn to the new argument, I will examine two versions of the indispensability argument which eschew holism, as the explanatory argument does, and fall right into the weasel's trap, like the versions of the argument of §3.

---

[11] See Burgess and Rosen 1997: 118.

[12] For present purposes, see Sober 1993 and Leng 2002: §3 and §6.

§5: Antecedents of the Explanatory Argument

Hillary Putnam formulated an early version of the indispensability argument which does not explicitly invoke holism. Call it the success argument.

PS
PS1. Mathematics succeeds as the language of science.
PS2. There must be a reason for the success of mathematics as the language of science.
PS3. No positions other than realism in mathematics provide a reason.
PSC. So, realism in mathematics must be correct.[13]

Putting aside, for a moment, worries about how realism explains the applicability of mathematics, PS is weak at its third premise. Any account of the applicability of mathematics to the empirical world other than the indispensabilist's would refute PS3. For example, Field 1980 argued that mathematics is successful as the language of science because it is conservative over nominalist versions of scientific theories. Mathematics is just a convenient shorthand for a theory which includes no mathematical axioms when cast in its most austere form.[14]

Michael Resnik's pragmatic indispensability argument similarly concludes that we are justified taking some mathematical statements to be true because of their utility facilitating scientific inferences.

RP
RP1. In stating its laws and conducting its derivations, science assumes the existence of many mathematical objects and the truth of much mathematics.
RP2. These assumptions are indispensable to the pursuit of science; moreover, many of the important conclusions drawn from and within science could not be drawn without taking mathematical claims to be true.
RP3. So we are justified in drawing conclusions from and within science only if we are justified in taking the mathematics used in science to be true.
RP4. We are justified in using science to explain and predict.
RP5. The only way we know of using science thus involves drawing conclusions from and within it.
RPC. So, by RP3, we are justified in taking mathematics to be true (Resnik 1997: 46-8).

---

[13] See Putnam 1975: 72-3.

[14] There are questions about whether mathematics really is conservative over nominalist versions of science. See Melia 1998 for some recent worries. MacBride 1999 is a useful survey.

Note that neither PS nor RP specify how one determines ontological commitments either within science or mathematics, leaving that matter elliptical. They both attempt to justify mathematical beliefs while avoiding worries about whether we must believe in the less-secure portions of scientific theories. Mathematics is clearly useful, or successful, as the language of science, Even if our best scientific theories are false, their undeniable practical utility still justifies our using them. Even if scientific theory turns out to be largely false, that falsity may not affect our beliefs in the mathematics we used to express that theory. A tool may work fine, even on a broken machine.

This flexibility comes at a cost. The inference to the truth of mathematic in RP1 is unjustified in the absence of a clear explanation of how science assumes the existence of objects. The utility of mathematics is not by itself an argument for its truth. We need a procedure for determining existence, or truth, or reference, to establish that these are to be found in science.

The same problem appears in RP2. The scientist may work without considering the question of mathematical truth at all: without employing a truth predicate applicable to mathematical statements, without taking mathematical theorems to be true. Again, we need a procedure for determining commitment, for knowing when the chips are down.

Against both Putnam and Resnik, the weasel, unchained by the abandonment of holism, attacks the strong conclusion that science presupposes or requires either the truth of mathematical claims or the existence of mathematical objects. A pragmatic argument for the indispensability of mathematics is no indispensability at all. All scientists need, whether we interpret their work as true or merely instrumentally useful, is the practical utility of mathematics. They need not presuppose mathematical truth.

For indispensability arguments which give up Quinean holism, instrumentalist interpretations of the mathematics used in scientific theory are compelling. Without holism and Quine's criterion for determining the ontological commitments of a theory, our existence claims need not be restricted to the quantifications of our best theory. We are free to adopt an eleatic principle, for example, as the fundamental criterion for existence. What makes QI immune to the weasel, and what makes the arguments from Putnam and Resnik liable to the weasel, is Quine's claim that the ontological chips are down precisely in our single best theory, and that we find our commitments in the quantifications of that holistic theory.[15] Putnam and Resnik tried to save the indispensability argument from problems arising from holism, but they opened the argument up to weaseling criticisms to which Quine's original argument was immune.

§6: The Explanatory Argument

According to the new explanatory indispensability argument, we should believe in mathematical objects because of their indispensable roles in our scientific explanations. Baker calls the argument the Enhanced Indispensability Argument.[16] We can see versions of the argument in Colyvan's recent work also. I'll follow Paolo Mancosu's formulation and call it the explanatory indispensability argument.

EI
EI1. There are genuinely mathematical explanations of empirical phenomena.
EI2. We ought to be committed to the theoretical posits postulated by such explanations.
EIC. We ought to be committed to the entities postulated by the mathematics in question (Mancosu 2008: §3.2).

The provenance of EI is a matter for dispute. Baker and Mancosu both misleadingly credit Field. "Hartry Field, one of the more influential recent nominalists, writes that the key issue in the platonism-nominalism debate is 'one special kind of indispensability argument: one involving indispensability for explanations' (Field 1989, p. 14)" (Baker 2005: 225).

Sorin Bangu follows Baker and Mancosu on this misdirection: "Field noted that even if, contrary

[15] Pincock 2004a: 62 makes a similar point.

[16] Baker 2009: 613. Baker originally presents the argument as a non-holistic alternative to QI; see Baker 2005: 224.

to what he argued in his (1980), mathematical posits turn out to be indispensable to scientific theorizing, they still can't be granted ontological rights until they are shown to be indispensable in a stronger, more specific sense; in particular, the realists should be able to show that mathematical posits are indispensable for scientific explanations (Field, 1989, pp. 14-20)" (Bangu 2008: 13-4).

A careful reading of the selection cited by both Baker and Bangu shows no such argument by Field. Field makes no claim that there is a heavier burden on the dispensabilist than recasting standard scientific theories to remove quantification over mathematical entities. Field's interest in explanation is exclusively on how explanatory merit factors into evaluations of our theories. His concern is with QI, where one factor in determining whether a theory is best is its explanatory force. Other factors include breadth and simplicity. For Field, once we have settled on a best theory, the only important question for the indispensabilist is whether that theory can be recast to avoid quantification over mathematical objects.

In the section cited, Field explicitly refers to his own work rewriting NGT, and he moves directly from talk about explanation to talk about theories.

> What we must do is make a bet on how best to achieve a satisfactory overall view of the place of mathematics in the world... My tentative bet is that we would do better to try to show that the explanatory role of mathematical entities is not what is superficially appears to be; and the most convincing way to do that would be to show that there are some fairly general strategies that can be employed to purge theories of all reference to mathematical entities (Field 1989: 18; see also fn 15 on p 20).

Field is clearly thinking of explanation on a metaphysical model, like a traditional covering-law account. In fact, Field says that an explanation is, "A relatively simple non-*ad hoc* body of principles from which [the phenomena] follow" (Field 1989: 15).[17]

---

[17] Mancosu notes that Field's discussion of the explanatory indispensability argument leads to questions about the success of his dispensabilist project, but does not make the point that if such projects were relevant to the success of the indispensability argument, then the argument in question must be the original version, and not the new explanatory argument.

The proponent of the explanatory argument puts aside the question of whether theories can be recast in order to eliminate mathematical entities and wonders whether non-mathematical explanations of physical phenomena are available. Indeed, recent work on EI grants the availability of nominalist reformulations of standard scientific theories and continues to urge that mathematical explanations of empirical phenomena support beliefs in mathematical objects. Exploring Malament's claim that phase-space theories resist dispensabilist constructions, Lyon and Colyvan write, "Even if nominalisation via [a dispensabilist construction] is possible, the resulting theory is likely to be less explanatory; there is explanatory power in phase-space formulations of theories, and this explanatory power does not seem recoverable in alternative formulations" (Lyon and Colyvan 2007: 242).

Given a metaphysical interpretation of 'explanation', Lyon and Colyvan's claim would be nearly nonsensical. Conserving metaphysical explanatory power is a standard requirement on nominalist reformulations, and it works unlike other theoretical virtues. One might wonder whether sacrifices in simplicity of ideology are worth parsimony in ontology or not. But scientists never give up metaphysical explanatory power in order to increase simplicity or parsimony. Indeed, Field constructs representation theorems precisely to support the claim that his reformulation lacks no metaphysical explanatory power of the standard theory. One just could not successfully nominalize a scientific theory by producing an alternative with less explanatory power unless one is using a non-metaphysical sense of the term.

Lyon and Colyvan's claim is plausible, though, if we interpret 'explanatory power' in the epistemic sense. Unlike standard scientific theories, dispensabilist reformulations will be imperspicuous, and not useful to working scientists. Dispensabilists generally do not say that scientists should adopt the reformulations. Indeed, Field grants that standard theories are more epistemically explanatory by arguing that mathematics is conservative over standard scientific theory, that adding mathematical axioms to nominalist theories will not allow one to derive any further nominalist conclusions. The nominalist wants to show that mathematics is conservative precisely because we are inevitably going to take advantage of the greater epistemic explanatory force of standard theories. Thus, the explanatory indispensability argument, in order to differentiate itself from QI, must rely on a notion of explanation that is not metaphysical.

We can also see that EI depends on an epistemic sense of explanation by noting that standard, metaphysical accounts of scientific explanation do not comfortably apply to mathematical explanation. Many mathematical inferences conform to standard criteria for scientific explanation, but are not explanatory. One can derive ‘2+2=4’ from basic axioms, but such derivations are not taken as explanations of the simple claim. In response, some philosophers of mathematics distinguish between explanatory and non-explanatory proofs. Another option would be to abandon the notion of mathematical explanation altogether.

A third alternative would be to claim that mathematical explanations of physical phenomena have two parts: a strictly mathematical explanation of a strictly mathematical theorem and a broader explanation of the physical phenomenon which invokes the mathematical theorem. The defender of the explanatory argument could claim that the nature of the pure mathematical explanation is isolated from, and thus irrelevant to, the nature of the broader explanation.[18] But unless the broader explanation were to take the mathematical result as a brute fact, the nature of the purely mathematical explanation is not isolatable. When we appeal to mathematical results in science, we want to know why they hold, if only to know their scope and limits. The proponent of EI, invoking mathematical explanation, is using the epistemic sense.

The distinction between whether the indispensabilist takes explanation to be a theoretical virtue, and thus is relying on QI, and whether the indispensabilist puts aside theoretical virtues and looks at the indispensability of mathematics for epistemic explanations, is subtle but important. The availability of a dispensabilist reformulation of a standard scientific theory is essential to QI. The availability of a

---

[18] Thanks to Alan Baker for raising this suggestion.

dispensabilist reformulation of a standard scientific theory is irrelevant to EI. A reformulation inevitably loses explanatory strength, in the epistemic sense.

There are two ways to view EI, in light of QI. We can see EI as an additional demand on the platonist, and thus an additional option for the nominalist. For example, Bangu and Melia say that even if dispensabilist constructions are not available, we should withhold commitments to mathematical objects since there are no genuinely mathematical explanations. On the Bangu/Melia view, the platonist has to show mathematics indispensable from both theories and explanations; the nominalist needs to show that mathematics is eliminable only from explanations or theories.[19]

In contrast, we can see the argument as an additional option for the platonist, and thus an additional demand on the dispensabilist. Baker and Lyon and Colyvan argue that even if the dispensabilist constructions are available, we should believe in mathematical objects as long as there are genuinely mathematical explanations of physical phenomena. On the Baker/Lyon and Colyvan view, the platonist needs to show that mathematics is indispensable only from explanations, and the nominalist must show how we can eliminate mathematics from both theories and explanations.

It will not matter here whether EI is taken as an additional burden on the nominalist or on the platonist. But, proponents and critics of EI might want to agree on an interpretation.

§7. Mathematical Explanations of Physical Phenomena

Debate over EI has focused on its first premise. Bangu, responding to Baker, argues that any purported mathematical explanation of a physical phenomenon is really a mathematical explanation of a mathematical phenomenon, and so question-begging against the platonist. I will first defend EI1 against Bangu's criticism, and then argue that such explanations do not decide the soundness of EI. The real

---

[19] See Bangu 2008, and Melia 1998: 70. Melia 2002 and Leng 2005: 179, though working with explanation as a theoretical virtue, can also be seen as taking this route.

problem is at EI2.

Examples of mathematical explanations of physical phenomena can be used to support either QI, taking explanatory merit as a theoretical virtue, or EI. Colyvan 2001 presents three examples intended mainly to support the claim that standard, mathematized theories have greater explanatory merit.

ME1. The bending of light. The best explanation of light bending around large objects is geometric, rather than causal.
ME2. Antipodes. The Borsuk-Ulam topological theorem, along with appropriate bridge principles, explains the existence of two antipodes in the Earth's atmosphere with the same pressure and temperature at the same time.
ME3. The Fitzgerald-Lorentz contraction. Minkowski's geometrical explanation of the contraction of a body in motion, relative to an inertial reference frame, relies on equations in four dimensions, representing the space-time manifold.[20]

Colyvan 2007 present three further illuminating examples.

ME4. Squaring the circle. That π is transcendent explains why we can not construct a square with the same area as a circle, using straight-edge and compass.
ME5. Mountaineering. A hiker, leaving base camp on one day and top camp the next, at the same time, will pass one point on the trail at the exact same time on both days.
ME6. Altruism. Simpson's paradox helps explain how a maladaptive trait like altruism can succeed despite the fact that altruistic populations, taken individually, are less fit.[21]

Not all of these examples are equally compellingly described as mathematical explanations of physical phenomena. Baker rightly worries about the status of the geometry on which ME1 and ME3 rely. If the relevant geometry is physical geometry, then the explanation may proceed without appeal to pure mathematics. Baker also argues that ME2 is a prediction rather than an explanation. There is no antecedent why-question since we are unlikely to discover the two antipodes, given both limitations on the precision of our instruments and independent interest in the phenomenon. Note that this complaint only holds against an epistemic account of explanation; one can easily provide a deduction or unifying

[20] See Colyvan 2001: 81-6.

[21] See Colyvan 2007: 120-1.

argument which yields the given phenomenon. Leng also complains that ME2 requires contentious idealizations, and so the requisite bridge principles will not apply. Still, these worries do not impugn the claim that if such antipodes were found on the Earth, the Borsuk-Ulam theorem would help explain them.[22]

Leng also calls ME4 a prediction. But, it is not a prediction of a physical fact. It is possible to construct a square with an area arbitrarily close to that of a given circle, by choosing arbitrarily close rational approximations of pi. We can draw a square-ish region with the same area as a given circle, within any given margin of measuring error. Still, if we had arbitrarily good measuring tools, we could always find a difference in the areas of the square and circle. If we found that we could not square the circle, the transcendence of pi might help explain that fact.

ME5 is more plausibly an explanation, rather than a prediction. But ME6 may not be true. Colyvan cites Malinas and Bigelow, but they only conclude that the mathematical result is worth examining since it *could* explain the persistence of altruism. "It is of considerable theoretical significance to explore the applications of Simpson's Paradox, to see whether this might help to explain not only the altruism but also the irrationality, inefficiency, laziness and other vices that may prevail in populations, and that can cause a population to fall short of the economic rationalist's or Darwinian's ideal of the ruthlessly efficient pursuit by each individual of its own profits or long-term reproductive success" (Malinas and Bigelow 2008).

Baker, interested in defending the claim that there are mathematical explanations of physical phenomena but worried about Colyvan's ME1-ME3, constructed an influential cicada example.

> ME7. Cicadas. That prime-numbered life-cycles minimize the intersection of cicada life-cycles with those of both predators and other species of cicadas explains why three species of cicadas of the genus Magicicada share a life cycle of either thirteen or seventeen years, depending on the environment.

[22] See Baker 2005: 226-7 and Leng 2005: 181-2.

Whether all of ME1-ME7, and any additional examples, work exactly as proponents of EI require is too strong a demand for establishing EI1. What is important is the underlying claim that there are mathematical explanations of physical phenomena.[23] If we take 'explanation' in the epistemic sense, as the defenders of EI should, that claim is overwhelmingly plausible. Baker and Colyvan's examples describe physical phenomena. They invoke mathematics to explain, in an epistemic sense, those phenomena. It may be possible to re-describe some of the phenomena or explanations either to eliminate or to isolate the mathematical elements. But, as they stand, such examples provide decisive, unsurprising evidence for EI1. Indeed, even the simple IM supports EI1. To defend EI1 from Bangu's criticisms, I will focus on ME7, which Bangu discusses and which is taken to be an especially strong example.

Three species of cicadas of the genus Magicicada share a life cycle of either thirteen or seventeen years, depending on the environment. The phenomenon of having prime-numbered life-cycles intrigued biologists, who sought an explanation. Baker claims that the phenomenon is explained thus:

CP
- CP1. Having a life-cycle period which minimizes intersection with other (nearby/lower) periods is evolutionarily advantageous.
- CP2. Prime periods minimize intersection (compared to non-prime periods).
- CP3. Hence organisms with periodic life-cycles are likely to evolve periods that are prime.
- CP4. Cicadas in ecosystem-type, E, are limited by biological constraints to periods from 14 to 18 years.
- CP5. Hence cicadas in ecosystem-type, E, are likely to evolve 17-year periods (Baker 2009: 614).

Baker argues that the mathematical explanans, at CP2, supports the empirical explanandum, at

[23] More examples: Lyon and Colyvan 2007 discuss how the honeycomb conjecture (a regular hexagonal grid represents the best way to divide a surface into regions of equal area with the least total perimeter) explains the structure of some bee hives. Mancosu 2008 notes that the twisting tennis racket theorem explains why a tennis racket, held horizontally and tossed to rotate about its intermediate principal axis, will make a near-half twist around its handle. Mancosu also cites Peter Lipton's observation that a simple geometric fact explains why a snapshot of a bundle of sticks tossed in the air will show significantly more of the sticks closer to the horizontal than to the vertical: there are more ways to be horizontal than vertical.

CP3. In fact, as Baker notes, CP3 is a "'mixed' biological/mathematical law." He uses this law to explain the further empirical claim CP5. Note that to apply the mathematical theorem used in CP2 to the case at hand at CP3, we need bridge laws, assurances that number theory applies to the cicadas' cycles. The pure number-theoretic premise refers to numbers and says nothing about life-cycles and their intersections.

Bangu argues that the explanandum in question at CP5 is, like CP3, a mixed statement composed of both mathematical and physical facts: a physical phenomenon (the time interval between successive occurrences of cicadas); the concept of a life-cycle period, expressed in years; the number seventeen; and the mathematical property of primeness. The mathematical explanation, he claims, only explains the mathematical portions of the explanandum. "Since primeness is a mathematical property, it is not surprising that we have to advance a mathematical explanation of its relevance, in terms of specific theorems about prime numbers" (Bangu 2008: 18).

Bangu is correct that assuming an explanandum with a mathematical element weakens the claim that there are mathematical explanations of physical phenomena. But as it stands, CP is an explanation of a biological fact which refers to mathematical objects. Bangu's allegation that the mathematical elements of CP1-4 only explain the mathematical portion of CP5 depends on whether we can reformulate CP5 to separate the mathematical portion from the empirical remainder of CP5. If the mathematical elements of CP5 were inseparable, then we could conclude, with the indispensabilist, that there are essentially mathematical elements of our descriptions of physical phenomena. The mathematical explanations of those elements will thus contribute indispensably to our explanations of the phenomenon. If the mathematical elements of the explanandum were truly ineliminable, then we would have reason to believe that the world is as the indispensabilist alleges. Bangu's claim that the mathematical elements of CP1-4 explain only the mathematical portion of the explanandum thus begs the question against the indispensabilist of whether the mathematical portion is essential to a description of the phenomenon.

Bangu's argument may seem plausible since the elementary use of numbers in CP5 is easily excisable. It is only drudgery to remove the reference to seventeen in CP5, no more sophisticated machinery is required than was used to rewrite IM as IN. The concept of primeness requires more machinery, though, as Leng observes, it does not even demand a completed ω-sequence.[24] Still, Baker rightly responds that if we eliminate the concept of primeness from the explanation, the application of this same general argument to ecosystems with other environmental constraints can seem coincidental.[25]

Bangu's criticism thus replays the dialectic between the indispensabilist and the dispensabilist. If we can construct explanations of physical phenomena which eliminate references to mathematical elements, then there are no essentially mathematical explanations of physical phenomena, and EI1 fails. If we can not reformulate our explanations to eliminate references to mathematical objects, then we have to grant EI1. But EI was introduced to avoid precisely this dispute.

Once we recast our explanations to remove references to mathematical objects, as the dispensabilist can try to do with all of ME1-7, we have traded satisfying (epistemic) explanations for appeals to austere, parsimonious theories. Moral 2 of §1 of this paper reappears: the theory we use to specify our ontological commitments may not be most useful when we want to explain (epistemically) facts about the world. If Bangu wants to defend nominalism in light of ME1-ME7, his claim should be either that we can reformulate such explanations to eliminate references to mathematical objects, taking a dispensabilist approach to explanations, or that the references to mathematical objects in the explanandum are not to be taken seriously, taking a weaseling approach.

Bangu's criticism of EI1 does not undermine the claim that there are mathematical explanations of physical phenomena, in the sense required by the proponent of EI. If we take 'explanation' in the metaphysical sense, Bangu's criticism seems promising, but only against the original QI. If we take

---

[24] See Leng 2005: 186.

[25] See Baker 2009: 617.

‘explanation’ in the epistemic sense, as the proponent of EI must, the claims ME1 - ME7 are thoroughly convincing.

§8. The Explanatory Argument: No Improvement on QI

Proponents of EI presume that their central challenge is to establish that there are mathematical explanations of physical phenomena. Even if not all of the examples work the way that Colyvan and Baker want them to work, taken together with the epistemic interpretation of ‘explanation’ they provide a compelling case for EI1. The real problem with such examples is that we want a reason to take such examples as expressing our ontological commitments. EI2, which claims that we ought to be committed to the objects postulated by mathematical explanations of physical phenomena, is problematic for familiar reasons: we need not be ontologically serious when we provide an epistemic explanation. Once we realize that the sense of ‘explanation’ in question is epistemic, any force that EI2 is supposed to have is lost. There is little reason to believe that explanations which facilitate our subjective understanding are ones in which we reveal our ontological commitments by speaking most soberly. EI seems plausible if we have a metaphysical sense of ‘explanation’ in mind; but then it’s no improvement on QI. EI is thus highly susceptible to weaseling.

Leng also argues that EI is susceptible to weaseling, but she derives her response to the explanatory argument from her response to QI. “If the original indispensability argument can be rejected on the grounds that some theoretical components can be good representations without being true (so that ‘fictional’ assumptions would do the representative work just as well), then the same considerations can be applied in the case of theoretical explanations” (Leng 2005: 187).

Leng relies on Melia’s claim that mathematics merely provides a language for representing or modeling physical facts, and that such representations need not be ontologically committing. “Nothing is lost in the explanation of cicada behavior if we drop the assumption that natural numbers exist” (Leng

2005: 186).

Leng also argues that weaseling responses are appropriate against ME2 and ME4. “We model the earth as a sphere, and pressure and temperature as continuous functions on the surface of this sphere. Once we have done this, the Borsuk-Ulam theorem can be seen to apply, and, to the extent that are [sic] model is a good one, we can draw a conclusion about the existence of a pair of points on the earth’s surface. Does the question of whether the sphere and the functions in our model *really exist* matter to the success of this piece of reasoning? It is hard to see how it should” (Leng 2005: 182).[26]

Against Leng and Melia, the aptness of a weaseling response to EI does not extend to QI. We can not expect our explanations to be the locus of our ontological commitments. But it is reasonable to expect our commitments to be represented by our best theories, in the sense required by QI. We constructed IN from IM precisely to be clear about our commitments.

Daly and Langford put the problem for the weasel attempting to deny the conjunction of QI2 and QI3, that we should believe in the objects in the domain of quantification of our best theory, in terms of a puzzle about rational belief. “How could it be rational to assert a theory if you believe both that the theory’s description of an abstract domain is false and that that description is indispensable in describing the concrete world?” (Daly and Langford 2010: 1115). At some point, we must speak seriously. But, since we should not expect our serious commitments to appear in our explanations, the liability of EI to weaseling does not extend to the original QI. Thus, the nominalist can not use my argument as a general strategy for resisting the indispensability argument.

Since EI is susceptible to weaseling, its defenders might re-cast the argument in a Quinean style, including explicit instructions for speaking seriously. Quine’s argument, re-cast for epistemic explanations, would say that our ontological commitments are to be found in our best explanations.

---

[26] Regarding ME4: “As with [ME2], the effectiveness of this explanation can be put down to the correctness of the model as a *representation* of the physical system, and not to the truth of the mathematics involved in this model” (Leng 2005: 184). Also see Leng 2005: 179, quoted in §3, above.

QEI
QEI1. We should believe explanations of our sense experience.
QEI2. If we believe explanations of our sense experience, we must believe in their ontological commitments.
QEI3. The ontological commitments of any explanation are the objects over which that theory first-order quantifies.
QEI4. The explanations of our sense experience first-order quantify over mathematical objects.
QEIC. We should believe that mathematical objects exist.

But QEI is no help at all to the defender of EI, since the conjunction of QEI1 and QEI2 is completely implausible. We need only believe our explanations in the sense in which we believe that we can re-cast them in ontologically serious ways. The chips are just not down in our epistemic explanations.

If the indispensabilist is tempted to believe in mathematical objects because of an explanation which uses mathematics, the explanation is not doing the work. The work is done by the background claim that there is a good theory supporting that explanation which requires those mathematical objects. Defenders of EI rely on a metaphysical notion of explanation in order to motivate the seriousness of our speech, and then switch to an epistemic notion of explanation in order to defend the viability of the claim that there are mathematical explanations of physical phenomena. If there are mathematical explanations of physical phenomena, and we want them to be taken as ontologically committing, then we have to find a way to fit the explanation into a traditional model so that QI applies. The explanatory indispensability argument is no improvement on QI.

§9. Peace in Our Time

We speak in ontologically serious tones only in the most austere version of our scientific theory. In a wide range of cases, explanations are better if they include references to mathematical objects. But since our explanations need not appeal to our most parsimonious theories, we need not take even the indispensable presence of mathematical objects in such explanations as ontologically serious.

Still, proponents of the explanatory indispensability argument present a serious case for the claim that there are mathematical explanations of physical phenomena. Even though that claim can not support the argument for which they use it, we might wonder about the importance of such explanations. There are two distinct attitudes that one can take. The weasel nominalist believes that their uses of mathematics give us no reason to believe in mathematical objects. The mathematics plays a representational role, perhaps akin to idealizations in physics. The indispensabilist can trot out compelling examples of applications of mathematics, but the weasel is really a mule, refusing to admit that any of these uses of mathematics are worth taking seriously.

The strength of the standard Quinean argument, which Field realized, but which more recent nominalists seem not to realize, is that we must, at some point, speak seriously. To Daly and Langford's claim that it is irrational to assert a theory without believing its description of mathematical objects, Melia responds that our expressive resources may be too impoverished to say what we want to say without invoking mathematics. "There is no a priori reason to suppose that the physical systems we wish to represent can *always* be characterized intrinsically using languages and theories that make no reference to the hypothesized abstract structures" (Melia 2010: 1119). But then, how do we determine the commitments of a theory?

The weasel returns to the pre-"Two Dogmas" point when philosophers questioned the existence of electrons because we couldn't see them. The strength of Quine's argument arises from realizing that we cannot invoke the physicist's theoretical commitments to electrons as reasons to believe in electrons without also being serious about the references to mathematical objects used in the same theories.

In his 1946 Harvard lecture on nominalism, Quine coins the term 'struthionism' to apply to those like Carnap, and Melia and Leng, who refuse to take references to mathematical objects within serious theories seriously. 'Struthionism' has the Greek word for ostrich at its core, so there's a third animal in the nominalist's menagerie: weasels, mules and ostriches.

I have defended these ragtag creatures in their claims that our explanations of physical phenomena need not impel us to believe in the referents of the mathematical terms they use. But there is still something uncomfortable about the attitude of the weasel, despite Melia's assurances that we can take back some portion of our serious assertions. Consider again Colyvan's mountaineering example. The explanation of the existence of a point on the mountain which the hiker passes at the same time on consecutive days includes a topological fixed point theorem. In siding with the weasel, I argued that such mathematical explanations give us no direct reason to believe in the existence of mathematical objects. We are still left wondering, though, whether we should believe in their existence or not, and whether we can learn anything from the fact that there are satisfying mathematical explanations of physical phenomena.

Explanations do seem to be more convincing if they do not refer ineliminably to fictional objects. If we had an alternate justification of our beliefs in mathematical objects, then explanations such as ME1 - ME7 would not appeal to fictional objects. It is not that the explanations themselves should be taken as ontologically serious. I am suggesting that an explanation which refers to fictional objects is less compelling than one which we can take fully literally.

Fortunately, we can maintain that the explanatory indispensability argument gives us no reason to believe in pure mathematical claims while also saying that our best explanations may use mathematical claims seriously. We merely need independent justifications of our beliefs in mathematical objects. Let's call such an independent route, if there were one, mathematical intuition. On the basis of intuition, we would believe (truly and *ex hypothesi* in a justified way), let's say, the axioms of ZFC. When we are doing science, we can thus appeal to mathematical machinery. But, we need not infer our mathematical knowledge from these uses. We're already justified in believing that mathematical objects exist.

The weasel rightly denies that we can justify our mathematical knowledge on the basis of our uses of mathematics in science. But she need not deny the existence of mathematical objects, and she need not give up an explanation of why mathematics is applicable to physical phenomena. Of course, most weasels are motivated by a deeper commitment to nominalism, and so will not find such a suggestion welcome. Moreover, the indispensabilist invokes QI and EI in order to try to convince the nominalist to believe in the existence of mathematics. So, neither the indispensabilist nor the weasel will find such a position attractive.[27] Still, independent considerations for or against the existence of mathematical objects are beside the point here. I am merely noting that the weasel's denial that the inference to mathematical objects from their indispensable uses in science is valid is compatible with the indispensabilist's claim that scientific explanations are more convincing if they refer to real objects.

§10: Conclusions

I presented two overlapping arguments. The first is that explanations which increase our understanding may not be those in which we express our ontological commitments. The second is that easy-road, or weasel, nominalism is an appropriate response only to some indispensability arguments, and not to Quine's original argument, which was designed to block such weaseling. Proponents of various alternate versions of the indispensability argument have weakened its ability to block the weasel.

In particular, the explanatory indispensability argument provides the platonist with no further ammunition than the traditional Quinean argument, QI. I did not defend QI. In fact, I believe that problems with Quine's holism and with the limitations on the platonism that the argument supports are serious. My claim is simply that QI is resistant to weaseling in ways that other arguments are not.

Lastly, I argued that EI is based on an epistemic concept of explanation. I contrasted the epistemic sense of 'explanation' with a metaphysical sense of the term on which QI is based. Perhaps there is an independent sense of 'explanation' on which EI might be based, and on which it would be more successful.

---

[27] Again, thanks to Alan Baker for stressing this point.

References

Azzouni, Jody. 2004. *Deflating Existential Consequence: A Case for Nominalism*. New York: Oxford University Press.

Baker, Alan. 2005. "Are there Genuine Mathematical Explanations of Physical Phenomena?" *Mind:* 114: 223-238.

Baker, Alan. 2003. "The Indispensability Argument and Multiple Foundations for Mathematics." *The Philosophical Quarterly* 53.210: 49-67.

Baker, Alan. 2001. "Mathematics, Indispensability and Scientific Practice." *Erkenntnis* 55: 85-116.

Balaguer, Mark. 2008. "Fictionalism in the Philosophy of Mathematics." *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/fictionalism-mathematics/>.

Balaguer, Mark. 1998. *Platonism and Anti-Platonism in Mathematics*. New York: Oxford University Press.

Bangu, Sorin Ioan. 2008. "Inference to the Best Explanation and Mathematical Realism." *Synthese* 160: 13-20.

Bueno, Otávio and Mark Colyvan. Forthcoming. "An Inferential Conception of the Application of Mathematics." *Noûs.*

Burgess, John. 1983. "Why I am Not a Nominalist." *Notre Dame Journal of Formal Logic* 24.1: 93-105.

Burgess, John, and Gideon Rosen. 1997. *A Subject with No Object*. New York: Oxford.

Carnap, Rudolph. 1950. "Empiricism, Semantics, and Ontology." In his *Meaning and Necessity: A Study in Semantics and Modal Logic*, second edition. The University of Chicago Press, 1956; Midway reprint, 1988.

Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.

Chihara, Charles. 1990. *Constructabiliy and Mathematical Existence.* Oxford: Oxford University Press.

Colyvan, Mark. 2010. "There's No Easy Road to Nominalism." *Mind* 119 (474): 285-306.

Colyvan, Mark. 2007. "Mathematical Recreation versus Mathematical Knowledge." In Leng, Mary, Alexander Paseau and Michael Potter, eds, *Mathematical Knowledge*, Oxford University Press, 109-122.

Colyvan, Mark. 2002. "Mathematics and Aesthetic Considerations in Science." *Mind* 111: 69-78.

Colyvan, Mark. 2001. *The Indispensability of Mathematics*. Oxford University Press.

Daly, Chris and Simon Langford. 2010. "Two Anti-Platonist Strategies." *Mind* 119: 1107-1116..

Field, Hartry. 1993. "The Conceptual Contingency of Mathematical Objects." *Mind* 102: 285-99.

Field, Hartry. 1989. *Realism, Mathematics, and Modality*. Oxford: Basil Blackwell.

Field, Hartry. 1980. *Science Without Numbers*. Princeton: Princeton University Press.

Friedman, Michael. 1974. "Explanation and Scientific Understanding." *The Journal of Philosophy* 71.1: 5-19.

Hellman, Geoffrey. 1989. *Mathematics Without Numbers*. New York: Oxford University Press.

Kitcher, Philip. 1981. "Explanatory Unification." *Philosophy of Science* 48.4: 507-531

Leng, Mary. 2005. "Mathematical Explanation." In Cellucci, Carlo and Donald Gillies eds, *Mathematical Reasoning and Heuristics*, King's College Publications, London, 167-189.

Leng, Mary. 2002. "What's Wrong With Indispensability? (Or, The Case for Recreational Mathematics)." *Synthese* 131: 395-417.

Lipton, Peter. 2004. "What Good is an Explanation." In John Cornwell, ed., *Explanations: Styles of Explanation in Science*, Oxford, pp 1-21.

Lyon, Aidan and Mark Colyvan. 2007. "The Explanatory Power of Phase Spaces." *Philosophia Mathematica* 16.2: 227-243.

MacBride, Fraser. 1999. "Listening to Fictions: A Study of Fieldian Nominalism." *The British Journal for the Philosophy of Science* 50.3: 431-55.

Maddy, Penelope. 1997. *Naturalism in Mathematics*. Oxford: Clarendon Press.

Maddy, Penelope. 1992. "Indispensability and Practice." *The Journal of Philosophy* 89: 275-289.

Malinas, Gary, and John Bigelow. 2008. "Simpson's Paradox." *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2008/entries/paradox-simpson/>.

Mancosu, Paolo. 2008. "Explanation in Mathematics." *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/mathematics-explanation/>.

Mancosu, Paolo. 2001. "Mathematical Explanation: Problems and Prospects." *Topoi* 20: 97-117.

Melia, Joseph. 2010. "Response to Daly and Langford." *Mind* 119: 1107-1116.

Melia, Joseph. 2002. "Response to Colyvan." *Mind* 111: 75-79.

Melia, Joseph. 2000. "Weaseling Away the Indispensability Argument." *Mind* 109: 455-479.

Melia, Joseph. 1998. "Field's Programme: Some Interference." *Analysis* 58.2: 63-71.

Pincock, Christopher. 2004b. "A New Perspective on the Problem of Applying Mathematics." *Philosophia Mathematica* 12.2: 135-161.

Pincock, Christopher. 2004a. "A Revealing Flaw in Colyvan's Indispensability Argument." *Philosophy of Science* 71: 61-79.

Potter, Michael. 2007. "What is the Problem of Mathematical Knowledge?" In *Mathematical Knowledge,* Mary Leng, Alexander Paseau, and Michael Potter, eds., Oxford University Press.

Putnam, Hilary. 1975. "What is Mathematical Truth?" In his *Mathematics, Matter, and Method: Philosophical Papers, Vol. I*. Cambridge: Cambridge University Press.

Putnam, Hilary. 1971. *Philosophy of Logic*. In his *Mathematics, Matter, and Method: Philosophical Papers, Vol. I*. Cambridge: Cambridge University Press.

Putnam, Hilary. 1962. "The Analytic and the Synthetic." In his *Mathematics, Matter, and Method: Philosophical Papers, Vol. II*. Cambridge: Cambridge University Press.

Quine, W.V. 1986. "Reply to Charles Parsons." In Hahn, Lewis Edwin and Paul Arthur Schilpp, eds. *The Philosophy of W.V. Quine*. La Salle: Open Court, 1986.

Quine, W.V. 1978. "Success and the Limits of Mathematization." In his *Theories and Things*. Cambridge: Harvard University Press, 1981.

Quine, W.V. 1960. *Word & Object*. Cambridge: The MIT Press.

Quine, W.V. 1958. "Speaking of Objects." In his *Ontological Relativity and Other Essays*. New York: Columbia, 1969.

Quine, W.V. 1955. "Posits and Reality." In his *The Ways of Paradox*. Cambridge: Harvard University Press, 1976.

Quine, W.V. 1951. "Two Dogmas of Empiricism." In his *From a Logical Point of View*. Cambridge: Harvard University Press, 1980.

Quine, W.V. 1948. "On What There Is." In his *From a Logical Point of View*. Cambridge: Harvard University Press, 1980.

Quine, W.V. 1946. "Nominalism." In Dean Zimmerman, ed., *Oxford Studies in Metaphysics*, Volume 4, Oxford University Press, 2004.

Quine, W.V. 1939. "A Logistical Approach to the Ontological Problem." In his *The Ways of Paradox*. Cambridge: Harvard University Press, 1976.

Railton, Peter. 1978. "A Deductive-Nomological Model of Probabilistic Explanation." *Philosophy of Science* 45: 206-226.

Resnik, Michael. 1997. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press.

Resnik, Michael. 1993. "A Naturalized Epistemology for a Platonist Mathematical Ontology." In, *Math Worlds: Philosophical and Social Studies of Mathematics and Mathematics Education*. Sal Restivo, et. al., eds. Albany: SUNY Press, 1993. Originally appeared in *Philosophica*.

Saatsi, Juha. 2007. "Living in Harmony: Nominalism and the Explanationist Argument for Realism." *International Studies in the Philosophy of Science* 21.1: 19-33.

Sober, Elliott. 1993. "Mathematics and Indispensability." *The Philosophical Review* 102: 35-57.

Van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Clarenden Press.

Van Fraassen, Bas. 1977. "The Pragmatics of Explanation." In Richard Boyd, Philip Gaspar and J.D. Trout, eds., *The Philosophy of Science*, Cambridge: The MIT Press, 1991.

Woodward, James. "Scientific Explanation." *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition), Edward N. Zalta (ed.), forthcoming URL = <http://plato.stanford.edu/archives/fall2009/entries/scientific-explanation/>.

Yablo, Stephen. 2005. "The Myth of the Seven." In *Fictionalism in Metaphysics*, M. Kalderon, ed., New York: Oxford University Press.

An Historical Introduction to the Philosophy of Mathematics

A Reader with Historical and Contemporary Selections

Proposal for Broadview Press

Russell Marcus
Department of Philosophy
Hamilton College
198 College Hill Road
Clinton NY 13323
rmarcus1@hamilton.edu
(315) 859-4056

and

Mark McEvoy
62-11 Woodside Avenue
Woodside, NY 11377
(516) 463 4991 (o)
(718) 651 5690 (h)

August 3, 2011

The enclosed table of contents proposes a reader in the philosophy of mathematics unlike any other available. We have designed it to be accessible to the undergraduate, providing the historical background required to grasp contemporary trends and a selection of contemporary position papers intended to provide glimpses of the current state of the discipline.

Despite the obvious narrowness of the field of philosophy of mathematics, this reader would be versatile. It would be broad enough to serve on its own as the basis for a first course in the philosophy of mathematics, for undergraduates or graduate students. It would also be useful as a supplement to other texts in a variety of courses. It might be used to supplement a more-sophisticated reader for advanced graduate courses in the philosophy of mathematics. It might be used as a part of courses in metaphysics or epistemology, focusing as it does on a critical area of debate in both of those disciplines. It would provide students with philosophical perspectives in formal courses in mathematical logic or set theory.

There are several good, recent introductory texts in the philosophy of mathematics, most notably James Robert Brown's *Philosophy of Mathematics: A Contemporary Introduction to the World of Proofs and Pictures* (Routledge); Stewart Shapiro's *Thinking About Mathematics* (Oxford); Alexander George and Daniel Velleman's *Philosophies of Mathematics* (Wiley-Blackwell), Daniel Bostock's *Philosophy of Mathematics: An Introduction* (Wiley-Blackwell); Michele Friend's *Introducing Philosophy of Mathematics* (Acumen). Our proposed volume would be excellent companions to any of these, providing primary sources for students reading those texts. Again, there is no historical reader accessible to undergraduates available.

The two venerable older collections in the philosophy of mathematics, Cambridge's expansive Benacerraf and Putnam, *Philosophy of Mathematics: Selected Readings* and W.D. Hart's *The Philosophy of Mathematics*, in the Oxford Readings in Philosophy series, are both a bit outdated and targeted to professional philosophers. Benacerraf and Putnam, while broad, has advanced selections which focus almost exclusively on developments in the early part of the last century, with no selections from the history of philosophy before Frege, and only a few papers dating after the 1960s. The Hart volume is also aimed at professionals, containing excellent, but exclusively-contemporary readings. Our reader, containing very little overlap with either volume, could easily be used in conjunction with either of these books.

There are several newer, advanced collections currently available: *The Oxford Handbook of Philosophy of Mathematics and Logic*, Shapiro, ed., is clearly aimed at professional philosophers, as are Jacquette's *Philosophy of Mathematics: An Anthology* (Blackwell) and Bueno and Linnebo's *New Waves in Philosophy of Mathematics* (Palgrave MacMillan). Our proposed introductory reader in the philosophy of mathematics will have a greater, because broader, appeal.

Popular books on mathematics appear to us to have a good market. The advanced readers in philosophy of mathematics are too sophisticated for that market. Our reader, being mainly historical, is more accessible and may also be interesting to the casual reader in philosophy and mathematics.

Philosophy of Mathematics courses are listed in the bulletins of almost all the universities listed in *The Philosophical Gourmet Report*. While these courses are admittedly not taught frequently to large groups, part of the reason for this is the lack of a friendly, accessible, historically-informative reader, the kind we are proposing. The field is experiencing a surge in popularity, with recent important articles in *Mind*, *Philosophy of Science*, *Synthese*, *Erkenntnis*, and *Philosophical Quarterly*; and a journal, *Philosophia Mathematica* dedicated exclusively to the area. The recently-published introductory texts in philosophy, especially the second edition of Brown; George and Velleman; and Bostock; demonstrate the growing market for just this sort of historical reader, accessible to undergraduates in both philosophy and mathematics.

We are sending this proposal to Broadview because of your commitment to precisely the market, undergraduates in philosophy, that we envision for our reader. We see this reader as a parallel to Peter Morton's wonderful *An Historical Introduction to the Philosophy of Mind*, which was, indeed, an inspiration for our collection.

The following table of contents is ambitious and likely demands some refinement. As it stands, it would be roughly 300,000 words without any editorial introductions. We also plan 30,000 - 50,000 words of original introductory material. Since we have been teaching from the primary sources, we have already made nearly all the selections indicated.

Table of Contents

Overview
- Benacerraf, "Mathematical Truth"
- Field, "Knowledge of Mathematical Entities"

Ancients
- Pythagoreans: Selections from Kline, *Mathematical Thought from Ancient to Modern Times*
- Plato, Selections from *Timaeus*; *Phaedo*; *Theaetetus*; *Republic*; *Meno*; and *Philebus*
- Aristotle, Selections from *Metaphysics* and *Physics*

Moderns
- Descartes, Meditation III; Meditation V; Selections from Second Replies
- Locke, Selections from the *Essay*
- Leibniz, "Meditations on Knowledge, Truth and Ideas"; Selections from *New Essays*
- Berkeley, Selections from *Principles*
- Hume, Selections from the *Enquiry* and the *Treatise*
- Kant, Selections from the *Critique* and the *Prolegomena*

19th Century
- Cantor, "On an Elementary Question in the Theory of Manifolds"
- Mill, Selections from *A System of Logic*
- Frege, Introduction to *Begriffsschrift*; Selections from *Grundlagen*

Early 20th Century
- Russell, Selections form *Problems of Philosophy*; Selections from *Introduction to Mathematical Philosophy*
- Hilbert and Bernays, *Foundations of Mathematics* (or Hilbert, "On the Infinite")
- Heyting, "Disputation"

Carnap, “Empiricism, Semantics and Ontology”
Ayer, Selections from *Language, Truth and Logic*
Wittgenstein, Selections from *Remarks on the Foundations of Mathematics*
Gödel, “What is Cantor’s Continuum Problem?”
Hintikka, Selections from *On Gödel* (or Smullyan, “The General Idea Behind Gödel’s Proof”)

Contemporary Views
Quine, “Two Dogmas of Empiricism”; “Existence and Quantification”; “On What There Is”
Chisolm, “Truths of Reason” (or Bonjour, *In Defense of Pure Reason*)
Benacerraf, “What Numbers Could Not Be”
Putnam, “What is Mathematical Truth?”

Field, Selections from *Science without Numbers* (or Introduction to *Realism, Mathematics and Modality*)
Kitcher, Selections from *The Nature of Mathematical Knowledge* (Or, “Mathematical Naturalism”)
Katz, Selections from *Cogitations, Metaphysics of Meaning*, and *Realistic Rationalism*
Balaguer, Selections from *Platonism and Anti-Platonism in Mathematics*
Resnik, “A Naturalist Epistemology for a Platonist Mathematical Ontology” (or Selections from *Mathematics as ths Science of Patterns)*
Burgess and Rosen, Selections from the Introduction to *A Subject with No Object*
Melia, Selections from “Weaseling Away the Indispensability Argument”
Baker, “Experimental Mathematics”
Putnam, “Why Nothing Works”

We could have a draft of the proposed volume complete by the end of next summer, 2012. We are flexible about suggestions and look forward to talking with you more about our proposal.

Thanks very much for your time and consideration.

readers would not be confused by this translation. Instead, if we alternatively take "zei" to mean "do harm to" and substitute it for "steal from" in the lines, the resulting sentences will surely make better sense.[4]

Another problematic translation is that of the sentence "shan yu ren tong, she ji cong ren," which was translated as "he was good at unifying himself with others. He put himself aside and joined with others" (p. 48). The subject of the passage is indeed on showing the admirable attitudes and manners by which virtuous and great people learn goodness from others. But Van Norden's translation here seems to have little to do with that, and this lack may well cause confusion to the reader. A more plausible reading is that "he (the Great Shun) regarded himself no different from others on (pursuing) good, and abandoned his own (evils) and followed others (on good)." The English words in the parentheses were possibly omitted in the original text. In 2A9.2, it would be better to translate the term "xian" as "talents," rather than "what is worthy" literally. Thus, "In taking office, he did not conceal his talents" is more readable than "In taking office, he did not conceal what is worthy" (p. 49).

In Van Norden's translation "and Qi and Chu were to attack it" in 3B5.1 (p. 80), the word "e" (hate) is missing. The correct translation would be "and Qi and Chu hated (its practice of benevolent government) and were to attack it." In the next passage, it is also questionable to translate "lao ruo kui shi" as "the young and weak offered the sacrificial food" (p. 80). It is not that the young and weak offered the sacrificial food, but that they offered food to the people of Bo who were sent to farm for Ge by King Tang. Indeed, the two expressions were put together and formed a complete sentence in the original text, which justifies why we read the expression this way.

Van Norden's second book, as its title suggests, contains those passages from the *Mengzi* that are taken to be the most important and essential ideas of Mengzi. By removing what he considered to be philosophically insignificant passages, this concise book offers the reader a chance of comprehending the core of Mengzi's philosophy more quickly. The selections, for the most part, are accurate in terms of their capturing what is essential in Mengzi's philosophy, but with several exceptions. One is that the translator selects 7A10 and abandons 7A21. Comparing the two, however, I do not see any reason why 7A21 is considered so less significant than 7A10 as to be not included. Actually, 7A21 contains important remarks on the relationships between *xing* (nature), *xin* (heart), and virtues, and should have been included in the selective translation. By the same token, 7B12, which addresses the practical importance of morality to the functioning of a state and society, should be no less important than 7B2 in Mengzi's philosophy. But 7B12 is not selected either.

In the second volume, the translation has been separated from the commentary. This change in format from the first book should be welcomed. The two formats together offer alternative ways of approaching Mengzi from which readers can select. If one likes to read the original text with the belief that reading commentaries would interfere and bias her potentially faithful understanding of it, she may choose Van Norden's second book. The spatial separation of the text from the commentaries gives her the opportunity to focus on the text alone. On the other hand, the first book may be preferred by those who find it convenient to make a quick reference to the helpful commentaries on the passages they are reading.

In conclusion, while the two books by Van Norden have several minor problems on his translation and commentaries,[5] they are truly successful and admirable. The works vividly show his meticulous research on Mengzi's philosophy. His insightful commentaries and notes, and the comprehensive introduction, shed light on the Mengzi's canon and are pedagogically invaluable to college students as well as the general reader who wants to study Mengzi's philosophy. Equally significantly, these two serious scholarly works represent important contributions made by him to the project of carrying the Mengzi's heritage into the western world.

Endnotes

1. Cordial thanks go to Dr. Tziporah Kasachkoff for her invitation to write this review. I appreciate the insightful comments of Dr. Eugene Kelly on an earlier version of the review, which identified many errors and helped to improve it significantly.
2. That justifies why many philosophy programs require their graduate students to comprehend at least one foreign language, so that they may read the original texts in the language in which they were originally written.
3. While Van Norden claims explicitly that his translation is often functional rather than literal in the preface, I found literal translations, though not necessarily nonfunctional ones, as I will show in what follows.
4. More on this. Mengzi always believes that benevolent governing by means of virtue is the correct and efficient way of keeping a state in order and prosperity, which would in turn strengthen one's kingship. By saying that a king is incapable of being virtuous, the subordinates actually deny the possibility and legitimacy of following a benevolent governing policy by the king, and thus in turn do harm to the latter. That is also why Mengzi tries to convince King Qi of his capabilities of being virtuous (1A7).
5. I pointed out these problems in this review for the purpose of alerting Prof. Van Norden, who might give a second thought to them if he publishes a second edition of the books.

The Ontological Argument from Descartes to Hegel

Kevin J. Harrelson (Amherst NY: Humanity Books, 2009), 255 pages, $39.98.

Reviewed by Russell Marcus
Hamilton College

The standard undergraduate modern philosophy survey course is an impossible monstrosity. The very idea of paying appropriate philosophical attention, in a mere fourteen weeks, to two extraordinarily fecund centuries of work on topics such as the relation of our minds to our bodies, the methods of science, the nature of space and time, free will and determinism, personal identity, justifications for civil society, and arguments for the existence of God is absurd. The Great-Figures solution to this absurdity limits one's syllabus to a few philosophers, say, Descartes, Hume, and Kant. The Great-Topics alternative covers a wider range of writers on a few themes such as substance, personal identity, and God. In addition to its undeniable utility as a research tool, Kevin J. Harrelson's new study, *The Ontological Argument from Descartes to Hegel*, could be a good addition to a Great-Topics course. The book covers a surprisingly wide range of modern writers, and could also be a useful text for an advanced course that focuses exclusively on the ontological argument.

Harrelson states serious critical goals for the book.

> I argue that the strategy for proving a priori the existence of God that remains in place during [the] period from Descartes' initial argument in the

> *Discourse on Method* (1637) to Hegel's final lectures in Berlin (1831), is both internally consistent and free of any easily identifiable error. More importantly, I try to show that the most common objections to the modern ontological proof…fail to identify any conclusive and universal fallacy. (18)

Harrelson divides the history of the ontological argument into three eras: pre-modern, including Anselm, Gaunilo, and Aquinas; modern, the focus of his volume; and post-Hegelian, which Harrelson mainly ignores. This division is a useful artifice, allowing Harrelson to focus on the era usually covered in Modern Philosophy courses. The book covers the standard presentations of the ontological argument (of Descartes, Spinoza, Leibniz, Kant, and Hegel), as well as expositions of it which are less well known (including those of More, Clarke, Wolff, Baumgarten, and Crusius). Harrelson's discussion of Huet's criticisms of Malebranche is amusing and useful, and his exposition of Mendelssohn's post-Kantian work on the argument is enlightening. His omission—with the exception of a few passing references—of Hume is curious. Despite the fact that Hume's criticisms of arguments for the existence of God generally focus on causal arguments (as in the *Dialogues Concerning Natural Religion*) the importance of Hume's principle that the truth-value of existence claims can never be discovered a priori deserves greater emphasis, especially in the discussion of Kant's work.

The inclusion of so many minor figures hinders Harrelson's narrative. In places, the book reads not as a monograph that traces the most important advances in the ontological argument but more like a dissertation, in which every mention of the argument by any minor figure is evaluated with every criticism taken to be worth remarking on. Still, Harrelson takes a firm critical stance toward the arguments.

Harrelson also impressively connects earlier work with later discussions of the argument. He consistently credits Aquinas for criticisms that might appear, to the student, as original with later writers. He connects Leibniz's work with that of Duns Scotus and Mersenne, and he cites Arnauld's anticipation of some of Kant's comments.

I would wager that a high proportion of philosophers, when prompted for the major flaw in the ontological argument, would point to Kant's claim that existence is not a predicate. Harrelson gives Gassendi proper credit for that point, and rightly notes that this point, standing alone, begs the question of whether God's existence is a single exception to the general rule that existence does not belong to the nature of an entity. Harrelson correctly insists that the full force and implication of the ontological argument cannot be understood when isolated from the specific contexts in which it appears, especially in the works of Malebranche, Spinoza, and Kant. Furthermore, Harrelson nicely shows that Kant's criticisms of the argument were aimed at versions of the argument found in the work of Leibniz, Wolff, and Baumgarten, and he argues, plausibly, that Kant was unfamiliar with the seventeenth-century expositions of the argument.

Dangers can arise from analyzing a short, if subtle, argument too finely. Harrelson divides Descartes's version of the ontological argument into what he deems its most thorough version, the syllogism from the First Replies, and what he calls the perfection argument. The First-Replies syllogism contains the premises that what we clearly and distinctly perceive as belonging to an object really does belong to that object, and that we clearly and distinctly perceive God's existence as belonging to his nature. The perfection argument alleges just that existence is a perfection. Harrelson follows Harry Wolfson in calling the First-Replies syllogism the primary Cartesian argument.

Aquinas had argued that linking the existence of a thing with its essence in one's thought need not entail that the thing exists independently of thought. The First-Replies syllogism alone does, as Harrelson says, serve to block this important Thomistic objection. Yet its minor premise remains completely unjustified without the addition of the perfection argument. Harrelson's division allows him to trace different portions of the argument through the subsequent two centuries, but only at the cost of losing track of the connections between them. The First-Replies syllogism is not plausible without at least the implicit assumption of the perfection argument. The perfection argument lacks any conclusion about the existence of God without the implicit assumption of the First-Replies syllogism, or something like it. Separating the two arguments is useful for tracing the history of the argument, but unfair for evaluating its success. Harrelson's fine distinction, while likely to be useful to historians of philosophy, will elude many undergraduates, creating more confusion than it merits, pedagogically.

Harrelson's exposition of Descartes's version of the argument might have benefitted from attention to the differences between Descartes's own goals for his analytic exposition, in the *Meditations*, and his synthetic exposition, in the Second Replies. I wonder if the difference between Descartes's presentation of the argument in the Fifth Meditation and his First-Replies syllogism can be explained more effectively by considering Descartes's distinctions between proof, demonstration, and explanation (on which see his Letter to Morin, 13 July 1638, AT II.197-8). Indeed, Harrelson could be a bit more sensitive to the difference between an argument and a proof; he sometimes calls the argument in question the "ontological proof."

In contrast, Harrelson neatly distinguishes versions of the argument which rely on intuitive awareness of God's existence from those which are intended as demonstrations or proofs. The book covers the role of Descartes's mathematical analogy (that existence belongs to God's essence the way that the sum of the angles of a triangle belong to the triangle), the question whether possible existence is attributable to a perfect being, and the worry that there is a gap in the argument between conclusions of the existence of a perfect being and that of a necessarily existent being. His discussion of different versions of the argument, such as those of Malebranche and Hegel, which minimize analogies from human existence to the existence of God and which conclude that being is, rather than that God exists, is helpful. These versions support Kant's "ontological argument" label against those who, finding the term misleading, prefer to call the argument the "a priori" argument, or the Cartesian argument. Harrelson also distinguishes versions of the argument aimed at combating atheism from those versions which would be compelling only to those who already believe in God's existence. I would have preferred less discussion of the latter, intuitive versions, which strike me as insufficiently philosophical.

Descartes's work provides a unifying theme for Harrelson's book. Still, the text would have benefitted from a concluding chapter, looking forward toward the post-Hegelian and contemporary proponents of the argument, especially since Harrelson calls the argument unassailable. Indeed, the lack of a unifying conclusion makes it difficult not to feel that Harrelson has failed to reach his stated goals, even though he has surveyed and criticized an admirable range of arguments and counter-arguments.

I enjoyed reading the book and learned from it, but I do not recommend it for classes in which instructors rely mainly on primary sources. The book does not include enough of the original source material for students to be able to grasp the critical commentary without also consulting the primary texts.

Also, while some chapters in the book stand on their own better than do others, most chapters refer indispensably to earlier discussions, so that students cannot profit from reading them in isolation of the discussions in earlier chapters to which they refer.

Nevertheless, I would recommend the book enthusiastically to students searching for paper topics. It could be valuable for a Great-Topics version of the standard course in modern philosophy, or for more advanced undergraduate and graduate classes covering seventeenth- or eighteenth-century metaphysics. Harrelson's study is accessible and nearly comprehensive over its target era. He generally avoids jargon. He helpfully names some of the major arguments, and provides a useful glossary for unfamiliar terms. Each chapter has many useful endnotes, and there is an excellent bibliography dividing the primary texts from the more recent secondary literature. The book contains a fine index.

I hope that publishers will encourage the production of similar manuscripts covering other salient topics in the modern era. A bookshelf full of such studies would be a valuable resource for the graduate student and beginning researcher. That Harrelson's text will be useful to undergraduates is an added bonus.

One small, final caveat: Harrelson's over-use of quotation marks is distracting, and sometimes misleading. For example, Harrelson writes:

> Descartes justifies this "predication rule" by appeal to the more general rule that "what is distinctly and clearly perceived is thereby true." (46)

The first set of quotation marks is otiose. While it is common to use quotation marks to indicate idiosyncratic usage, Harrelson uses them in almost every paragraph of the book, often repeatedly even within a single sentence. The words contained in the second set are a paraphrase, not a quote, of the cited section. The sentence would be better rendered without any quotation marks at all. Such infelicities are especially unfortunate since the ontological argument requires careful distinctions between uses and mentions, between concepts and objects, and between thoughts and concepts. I would not recommend the text to an undergraduate without first discussing proper usage.[1]

Endnotes

1. Thanks to Shoshana Brassfield for helpful comments.

List of Books Received

Baier, Annette C. *The Cautious Jealous Virtue: Hume on Justice* (Cambridge, MA: Harvard University Press, 2010).

Cohen, Jonathan R. *Science, Culture and Free Spirits. A Study of Nietzsche's* Human, All Too Human (Amherst, NY: Humanity Books, 2010).

Gale, Richard M. *John Dewey's Quest for Unity: The Journey of a Promethean Mystic* (Amherst, NY: Prometheus Books, 2010).

Gordon, Sarah Barringer. *The Spirit of the Law. Religious Voices and the Constitution in Modern America* (Cambridge, MA: Belcamp Press of Harvard University Press, 2010).

Halliday, Paul D. *Habeus Corpus. From England to Empire* (Cambridge, MA: Belcamp Press of Harvard University Press, 2010).

Mirel, Jeffrey E. *Patriotic Pluralism. Americanization Education and European Immigrants* (Cambridge, MA: Harvard University Press, 2010).

Mizzoni, John. *Ethics: The Basics* (Malden, MA: Wiley-Blackwell, 2010).

Shook, John R., ed. *Exuberant Skepticism: Paul Kurtz* (Amherst, NY: Prometheus Books, 2010).

Sternberg, Eliezer J. *My Brain Made Me Do It: The Rise of Neuroscience and the Threat to Moral Responsibility* (Amherst, NY: Prometheus Books, 2012).

Warner, Michael, Jonathan VanAntwerpen, and Craig Calhoun, eds. *Varieties of Secularism in a Secular Age* (Cambridge, MA: Harvard University Press, 2010).

Warren, James, ed. *The Cambridge Companion to Epicureanism* (New York: Cambridge University Press, 2009).

Addresses of Contributors

John Immerwahr
Department of Philosophy
Villanova University
Villanova, PA 19010
www.homepage.villanova.edu/john.immerwahr
john_immerwahr@villanova.edu

Russell Marcus
Department of Philosophy
Hamilton College
198 College Hill Road
Clinton, NY 13323
rmarcus1@hamilton.edu

Xuetai Qi
School of Humanities
Nanjing University of Posts & Telecommunications
9 Wenyuan Road
Nanjing 210046, China
xuetaiqi@gmail.com

Dr. Jonas Pfister
Institut für Philosophie
University of Bern
Länggass-Str. 49a
3000 Bern 9
Switzerland
Jonas.pfister@bluewin.ch

A Cooperative-Learning Lesson Using the *Objections and Replies*

Published in *The APA Newsletter on Teaching Philosophy* 9.2, Spring 2010: 2-5.

Russell Marcus

Chauncey Truax Post-Doctoral Fellow

Department of Philosophy, Hamilton College

198 College Hill Road

Clinton NY 13323

rmarcus1@hamilton.edu

(315) 859-4056 (office)

(315) 381-3125 (home)

(315) 335-4711 (mobile)

October 11, 2009

3705 words

Abstract:

Descartes's *Meditations* are a staple of philosophy courses at all levels. The *Objections and Replies* to the *Meditations* are both an integral part of the original work and a fecund resource for philosophy teachers. In this article, I explain how I use excerpts from the *Objections and Replies* for an in-class cooperative-learning exercise.

Introduction

Descartes's *Meditations* are a staple of philosophy courses at all levels. The *Objections and Replies* to the *Meditations* are both an integral part of the original work and a fecund resource for philosophy teachers. In this article, I explain how I use excerpts from the *Objections and Replies* for an in-class cooperative-learning exercise. I begin by discussing the relevance and utility of the material I use in the lesson.

Historical Background

The first edition of Descartes's *Meditations on First Philosophy* was published in Paris in 1641. Descartes was concerned that his work be shown defensible in the face of thorough criticism, and that he deflect attacks. "I would have liked to have the approbation of a number of people so as to prevent the cavils of ignorant contradiction-mongers."[1] To this end, he included not just the six brief Meditations and introductory notes, but also a much longer collection of six sets of objections from theologians, scholars, and friends, as well as Descartes's replies to these objections. The objectors were:

1. Johan de Kater (Caterus), a Catholic Dutch theologian;
2. Various theologians and philosophers in a circle centered around the friar and mathematician Marin Mersenne;[2]
3. Thomas Hobbes, in his 50s, exiled and living in France, still ten years before the publication of *Leviathan,* and a year before *De Cive*;
4. Antoine Arnauld, philosopher and Jansenist theologian, a co-author of *The Port-Royal Grammar*, whose comments Descartes said he preferred;[3]
5. Pierre Gassendi, French atomist philosopher; and
6. Various theologians and philosophers whose comments were again collected by Mersenne.

The second edition of the *Meditations*, published a year later in Amsterdam, included an additional, harsh seventh set of objections from Pierre Bourdin, a Jesuit priest, along with Descartes's replies.

Descartes had good reason for concern about the acceptability of his work. Galileo's condemnation by the Inquisition in June 1633 created a dangerous climate for Descartes, who was just entering his most productive philosophical period. Descartes immediately scrapped his plans to publish *Le Monde*, which presented a heliocentric system, as well as the foundations of physics and human physiology. In 1637, when Descartes resolved to publish essays on optics, geometry, and meteorology (though omitting the most controversial topics), he did so anonymously. The introductory essay, now known as the *Discourse on Method*, nonetheless provoked severe criticism. Descartes prepared for publication his consequent correspondence with Jean-Baptiste Morin, a professor at the Collège de France who later contributed indispensably to the second set of objections. Morin and Descartes, though, abandoned their plan when it became clear that their differences on many details could not be resolved.

The *Objections and Replies* to Descartes's *Meditations* are no mere auxiliary commentary to a more important, central work. Indeed, they are essential to the *Meditations* themselves. "[I]t would be illegitimate to read the *Meditations* in abstraction from the *Objections and Replies* with which they intentionally form an organic whole..." (Marion 1995: 20). Descartes scholars, of course, know the most important elements of the *Objections and Replies*, as well as the core text.

Still, students studying Descartes for the first or second time, in introductory courses, or in undergraduate surveys of modern philosophy, tend to see only very little, if any, of the *Objections and Replies*. Introductory philosophy readers and texts almost always include selections from the *Meditations*, if not the complete text. Yet, few of the standard introductory readers include selections from the *Objections and Replies*. Many popular editions of the *Meditations* include the *Discourse on Method*, but not the longer *Objections and Replies*.[4] Even the best modern philosophy readers include only a few, sample, objections.

This widespread neglect by philosophy teachers of the greater portion of the original work is understandable. Teachers in both introductory and modern philosophy courses are often too rushed to spend more time on Descartes. The larger themes (e.g. the disputes with Gassendi on empiricism and atomism) can overwhelm the undergraduate. Marion's claim that it is illegitimate to read the *Meditations* without studying the *Objections and Replies* is too strong when applied to the teaching of undergraduates. The *Meditations* may stand on their own in our classrooms.

But Descartes's oeuvre is not the product of a solitary meditator, working alone. His exchanges with colleagues are edifying. The *Objections and Replies* were passed among the objectors sequentially, so themes can be traced through these exchanges. Caterus, Arnauld, Gassendi, and Mersenne all criticize the arguments for the existence of God. Mersenne, Hobbes, and Gassendi work on the criteria of clear and distinct ideas and the problem of Cartesian circularity. All of the objectors comment on the mind/body distinction. Descartes and his objectors repeatedly pursue and elaborate arguments first raised earlier in the *Objections and Replies*. For example, after Gassendi uses the example of a straight stick appearing bent in water to raise a worry about Descartes's account of error (AT VII.333), Mersenne returns to the example to argue that the senses, rather than reason, correct the error (AT VII.418). While scholars are able to tease out the interwoven thematic threads of conversation, the student may well find it impossible to follow a single train of argument across different objectors.

The *Objections and Replies* can be alluring to undergraduates, if presented appropriately. Witty and acerbic, they include lively and memorable examples that bring the *Meditations* into sharper focus. The lesson I describe presents small, thematically-organized portions of the objections, across objectors. I have refined the following lesson over several years, and the students generally respond favorably.

The Lesson

I introduce my students to the *Objections and Replies* by using a cooperative-learning exercise in which students in three-membered groups adjudicate three objections and replies. This lesson is best suited to a class of at least an hour and fifteen minutes. I generally place the lesson at the end of the portion of a course in which we have been reading and discussing the *Meditations*, before the students have taken an exam or written a paper.

The materials needed for this exercise are just selections of objections and replies for the students to adjudicate. I prepare sets of objections and replies, gently edited for brevity and focus, and organized topically so that each group can focus on one theme. Another option would be to assign pre-determined selections from an assigned text. The second, third, and sixth sets of objections easily lend themselves to the exercise. Cottingham, Stoothoff, and Murdoch's single-volume collection of Descartes's work contains a sufficient number of objections for a small group exercise. The full set of objections and replies is available in the second volume of their three-volume collection. The Ariew and Cress collection is less expensive and presents a useful selection.[5] I have not asked my students to prepare for the class by reading selections in advance, though one could do so. I prefer that the students discover the assigned material together, in their groups.

Good group assignments are essential to the success of any cooperative lesson. For this lesson, I prefer random group assignments.[6] The easiest method for random group assignments involves counting-off by the number of students in the class divided by the number of people in each group; all the ones form a group, all the twos, all the threes, etc. For the *Objections and Replies* lesson, I use a more interesting way of forming groups. I have printed and laminated pictures of Descartes and his objectors. I cut some of these pictures into three pieces, and two of them into four pieces (for classes with a number of students not divisible by three). In class, I mix the pieces into a basket which the students pass around, each drawing a piece, while I introduce the lesson. Groups are formed by the students finding the puzzle pieces that fit with their own. Students find their groups quickly, and they giggle a bit about the technique.[7]

Once groups are formed, each group either chooses a topic or is assigned one. Different groups can work on the same theme, or each group can focus on a different theme. In introductory courses, I prefer to let students choose topics that interest them. In smaller classes, I have laid six to eight piles of copies of my edited selections on a side table, with topic names prominently displayed, and allowed the students briefly to shop for a topic. In more advanced courses, a desire to cover a range of specific topics may take precedence, and the instructor may wish to assign specific themes to different groups. Once the readings for each group are distributed, the adjudications can begin.

There are three roles within each group, and three adjudications, so that each member of the group can play each role once. The roles are: 1. Objector; 2. Descartes; and 3. Facilitator/Scribe. The Objector reads the objection aloud to the group. Descartes reads Descartes's reply. Then, all the students in the group discuss the merits of the objection and reply. The Facilitator adjudicates and takes notes. During adjudication, the Objector and Descartes lobby the Facilitator and defend their positions. I urge all members of the group to seek agreement on a result. Thus, to help adjudicate, the students must both play their roles, and step out of them.

The Facilitator/Scribe transcribes a summary of the group's verdict. Not all adjudications result in a clear victor. Many of the debates are best read as referring to larger disagreements. Some exchanges introduce new concepts that require more serious reflection. For example, in response to Arnauld's objection that the argument for the distinction between mind and body would also lead to a distinction between a right triangle and figure whose side lengths are Pythagorean triples, Descartes refers to adequate and complete knowledge. These terms do not appear in the *Meditations* proper, and are typically new to my students. A group adjudicating this objection may decide that further research into adequate and complete knowledge is required to achieve a verdict. Especially in such cases, the scribe should write down questions for further research. I ask all members of each group to be prepared to present at least one of their three adjudications to the whole class at the end of the group work.

Once the first objection is adjudicated, the students within each group switch roles for the second and third adjudications. If there are one or two groups of four, some members of those groups may not play each role. In such cases, the facilitator and scribe roles can be separated. Alternatively, such groups can be assigned four objections and replies.

After each group has adjudicated each objection and reply, groups dissolve and the class comes together for discussion. In longer classes, I ask individual students to present one of their results, including summaries of an objection and reply and the group's adjudication. Instructors interested in enforcing individual accountability can easily select students at random to present their results.[8] At the end of the class, I normally collect the groups' transcriptions. I read them and discuss the exercise and some of the results in a future meeting. For classes needing extra motivation, the transcripts may be graded.

Benefits of the Lesson

Reading excerpts from the *Objections and Replies* has allowed my students not only to understand the *Meditations* better, but also to see connections among both earlier and later philosophers. Many lasting criticisms of Descartes's work are initially voiced in the *Objections and Replies*. For example, Mersenne, in the second set of objections, immediately raises the problem of Cartesian circularity. Arnauld objects to innate ideas by considering the ideas of children and madmen. Gassendi insists that existence is not a perfection and claims that the cogito presumes that whatever thinks exists. Studying the *Objections and Replies* allows students first-hand access to Descartes's own responses to these well-known difficulties.

Perhaps more importantly, reading the *Objections and Replies* has provided my students with insights into philosophical methodology. When my classes start working on the *Meditations*, I try to present Descartes's work charitably, seeking understanding and a good interpretation. When we get to the *Objections and Replies*, the students engage the text differently. They are no longer attempting merely to figure out how Descartes's work is best understood. They are now trying to figure out whether he is right. The exchanges in the *Objections and Replies* model actual philosophical discourse, including a wide range in quality of arguments, and Descartes does not always have the better ones. No longer the wise master, Descartes is one among a group of peers, all of whom have their teeth sunk deeply into puzzling questions, and are actively engaged with one another, and very human. Descartes's defensiveness emerges, especially in his responses to Hobbes and Gassendi. Gassendi's arguments against Descartes's mind/body distinction, for example, are deeper than Descartes's dismissive tone toward him indicates. By studying these exchanges, students can learn that they must take their critics seriously, rather than dismissing or insulting them.[9]

While learning about philosophical methods is valuable, some students can become confused by the critical freedom the *Objections and Replies* exemplify. The shift from analyzing the *Meditations* to adjudicating the *Objections and Replies* can disorient the unprepared student at the same time as it empowers the stronger young philosopher. This lesson actually makes students think, and they recognize it. As I walk around the classroom during adjudications, I have to assure the students that they are free to evaluate the arguments themselves. They can see that I am just as puzzled as they are about some of the arguments. The lesson gives me an opportunity to show, emphatically, that my authority as their teacher does not derive from my knowing all the answers.

In the remainder of this section, as a last illustration of the benefits of this cooperative lesson, I discuss some sample content. For brevity's sake, I will consider only a few objections and replies on one theme: the argument for the distinction between the mind and the body.[10] Let's take Descartes's original argument to be as follows:

1. Whatever I can clearly and distinctly conceive of as separate, can be separated by God, and so are really distinct.
2. I have a clear and distinct understanding of my mind, independent of my body.
3. I have a clear and distinct understanding of my body, independent of my mind.

So, my mind is distinct from my body

Caterus, citing Duns Scotus, complains about the major premise that the ability to conceive the mind as separate from the body does not entail that there is a real distinction between them. "The formal concepts of the two are distinct prior to any operation of the intellect, so that one is not the same as the other. Yet it does not follow that because [God's] justice and mercy can be conceived apart from one another they can therefore exist apart" (AT VII.100).

In response Descartes concedes that we can separate, in our minds, things that can not really be separated: motion and shape from a body, justice or mercy from just or merciful people. But, these sorts of abstractions only apply to incomplete entities.

> [The formal distinction] applies only to incomplete entities. By contrast, I have a complete understanding of what a body is when I think that it is merely something having extension, shape and motion. I deny that it has anything which belongs to the nature of a mind. Conversely, I understand the mind to be a complete thing, which doubts, understands, wills and so on, even though I deny that it has any of the attributes which are contained in the idea of a body. This would be quite impossible if there were not a real distinction between the mind and the body (AT VII.120-1).

Arnauld picks up on the contrast between complete ideas of a thing and clear and distinct ideas. He argues that we can clearly and distinctly understand that a triangle is right-angled while failing to understand that the Pythagorean theorem holds of that triangle. Similarly, our clear and distinct understanding of our minds could be compatible with materialism. In response, Descartes emphasizes the importance of the minor premises of the original argument.

> Although we can clearly and distinctly understand that a triangle in a semi-circle is right-angled without being aware that the square on the hypotenuse is equal to the squares on the other two sides, we cannot have a clear understanding of a triangle having the square on its hypotenuse equal to the squares on the other sides without at the same time being aware that it is right-angled. And yet we can clearly and distinctly perceive the mind without the body *and* the body without the mind (AT VII.224-5, emphasis added).

Note the opportunities for discussion here: Is Descartes correct that we can not understand that the Pythagorean theorem holds of a polygon without understanding that it has a right angle? Does the original argument work with only one minor premise? Is our understanding of our mind or our body really complete? Mersenne's objections also raise these worries. "What if [the thinking thing] turned out to be a body which, by its various motions and encounters, produces what we call thought? Although you think you have ruled out every kind of body, you could have been mistaken here, since you did not exclude yourself, and you may be a body. How do you demonstrate that a body is incapable of thinking, or that corporeal motions are not in fact thought?" (AT VII.122-3; see also Hobbes's related concerns at

AT VII.172-4).

Lastly on the mind/body distinction, these selections are useful for helping students to avoid taking Descartes's arguments to establish that the body is inessential to the self. Arnauld makes the worry explicit. "It seems that the argument proves too much, and takes us back to the Platonic view (which you reject) that nothing corporeal belongs to our essence, so that man is merely a rational soul and the body merely a vehicle for the soul, a view which gives rise to the definition of man as a soul which makes use of a body" (AT VII.203).

Again, Descartes's use of concrete examples expands and illustrates the original, all-too-brief discussion in the *Meditations*. "[S]omeone who says that a man's arm is a substance that is really distinct from the rest of his body does not thereby deny that the arm belongs to the nature of the whole man. And saying that the arm belongs to the nature of the whole man does not give rise to the suspicion that it cannot subsist in its own right" (AT VII.228).

Summary

Well-constructed cooperative-learning exercises may be distinguished from simple group work by attention to four factors: 1. Careful distribution of students into groups; 2. Assignments of specific roles and responsibilities to each member of the group; 3. Specific and attainable objectives; and 4. A balance of emphasis on both group dynamic and individual accountability.[11] In this lesson, I carefully distribute students into small, random groupings. Group members have specific duties throughout the lesson. The students have clear, specific goals at all times: adjudicating all three assigned objections and replies, and making a list of questions for further research. The groups are charged with preparing all students to present one of their adjudications, and individual accountability is easily enforced.

I love working with the students on this lesson. It gives them the tools to discuss philosophy with each other in precisely the ways that I want them to continue outside of class. When the lesson comes, as I use it, at the end of an extensive treatment of the *Meditations*, most of the students have learned enough content to engage the exchanges confidently, and to have fun with it.

In addition to this enjoyable lesson, exposing students to the topics in the *Objections and Replies* is an excellent way to initiate research for essays. Thus, the exercise has the added benefit of providing the instructor with a wide range of new paper topics, helping to alleviate the tedium of reading the same themes term after term.

Besides its obvious application in modern philosophy courses, I have used this in-class exercise successfully in introductory courses, and, in abbreviated fashion, in a sophomore-level philosophy of mind course. If the particular objections and replies are chosen appropriately, they require no prior philosophical background, other than the *Meditations* themselves, and so are a good supplement to any course in which the *Meditations* are read.[12]

Works Cited

Ariew, Roger and Donald Cress, eds. 2006. *Descartes: Meditations, Objections, and Replies*. Indianapolis: Hackett Publishing Company.

Ariew, Roger and Marjorie Grene. 1995. *Descartes and His Contemporaries: Meditations Objections, and Replies*. Chicago: The University of Chicago Press.

Cottingham, John, Robert Stoothoff, and Dugald Murdoch, eds. 1984. *The Philosophical Writings of Descartes*, Volumes I-II. Cambridge University Press.

Cottingham, John, Robert Stoothoff, and Dugald Murdoch, eds. 1988. *Descartes: Selected Philosophical Writings*. Cambridge University Press.

Cottingham, John, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny, eds. 1991. *The Philosophical Writings of Descartes*, Volume III: The Correspondence. Cambridge University Press.

Garber, Daniel. 1995. "J.-B. Morin and the *Second Objections*." In Ariew and Grene 1995.

Johnson, David W., Roger T. Johnson and Karl A. Smith. 1998. *Active Learning: Cooperation in the College Classroom*. Edina: Interaction Book Company.

Kagan, Spencer. 1992. *Cooperative Learning*. San Juan Capistrano, CA: Kagan Cooperative Learning.

Marcus, Russell. In preparation, A. "Comments on Cooperative-Learning Group Assignments."

Marcus, Russell. In preparation, B. *Themes in the Objections and Replies: Selected Objections and Replies to Descartes's* Meditations, *Organized Topically with New Introductory Material*

Marion, Jean-Luc. 1995. "The Place of the *Objections* in the Development of Cartesian Metaphysics." In Ariew and Grene 1995.

Slavin, Robert. 1995. *Cooperative Learning: Theory and Practice*, 2nd edition. Boston: Allyn & Bacon.

Endnotes

1. Letter to Mersenne, 30 September 1640, AT III.184. Translation from Cottingham et al., vol. 3.

2. The second set of objections was collected and presented by Mersenne, who handled the remaining circulation of the manuscript. Cottingham attributes the second set directly to Mersenne (Cottingham 1984 v. II: 64). Garber 1995 attributes many of those objections to Jean-Baptiste Morin.

3. "I consider them the best of all" (Descartes, letter to Mersenne, 4 March 1641, AT III.331).

4. Hackett's *Meditations, Objections, and Replies*, edited by Ariew and Cress, contains much of the *Objections and Replies*, though omitting most of the extended exchange with Gassendi and all of the long seventh set.

5. I am currently preparing a ms (Marcus, In preparation B) with a complete selection, parsed into over 100 exchanges, arranged topically.

6. Elsewhere, I argue for this preference. See Marcus, In preparation, A.

7. See Johnson, Johnson, and Smith 1998: 2-8 for other methods of randomly assigning students to groups.

8. I have not yet had the opportunity and inclination to ask the students to summarize their findings in a class wiki, though such an activity, outside of class, seems well-suited to the lesson.

9. Thanks to an anonymous referee for helping me formulate this paragraph, and the next.

10. I would be pleased to email interested readers further examples of the benefits of reading the *Objections and Replies*.

11. Slavin 1995 reviews a variety of cooperative-learning techniques, and presents a useful survey of classroom research on its effects.

12. Thanks to four anonymous referees for helpful and encouraging comments, and to the participants in my session at the August 2008 meeting of the American Association of Philosophy Teachers in Guelph.

Embracing the Cartesian Circle
January 4, 2012

Work in Progress

I. The problem of circularity

Behold the problem of the Cartesian circle: In the First Meditation, Descartes presents three arguments for doubt, the strongest of which appears to call into question all of his beliefs. In the Third Meditation, attempting to reconstruct his beliefs, Descartes presents a causal argument for the existence of God which depends for its legitimacy (in overcoming the doubts) on the premise that whatever is clearly and distinctly perceived is true. In the Fourth Meditation, Descartes argues that whatever is clearly and distinctly perceived is true, relying on the premise that God exists, and is perfectly good.[1] "The central argument of Descartes's metaphysics proceeds through clear and distinct perception to the existence of God, and then from God's veracity to the truth of clear and distinct perceptions" (Gewirth 1941: 368). Descartes appears to have reasoned in a circle.

The problem of the Cartesian circle has been noted early and often. The Second Objectors had questioned (AT 126-7) whether the criteria of clarity and distinctness would secure the beliefs that Descartes hoped it would. But Arnauld is ordinarily credited with the first precise statement of the circle.

> I have one further worry, namely how you avoid reasoning in a circle when you say that we are sure that what we clearly and distinctly perceive is true only because God exists. But we can be sure that God exists only because we clearly and distinctly perceive this. Hence, before we can be sure that God exists, we ought to be able to be sure that whatever we perceive clearly and evidently is true (*Fourth Objections*, AT VII.214).[2]

One natural response to the problem, especially for contemporary readers with a bit of logical acumen and a distaste for arguments for God's existence, is to dismiss whatever depends on either the criteria of clarity and distinctness or on the causal argument in the *Meditations*. Thus, many of us interpret Descartes as stuck in a solipsistic rut, establishing nothing much beyond his own existence. Indeed, many introductory philosophy and epistemology collections present only the first two Meditations, generating Descartes's skeptical worries, and then ignoring Descartes's positive program beyond the cogito.

In contrast, several considerations compel us to take the rest of Descartes's positive project seriously. First, it is worthwhile to note that Descartes did not think that there was a problem with the circle.

> I was not guilty of circularity... We are sure that God exists because we attend to the arguments which prove this; but subsequently it is enough for us to remember that we perceived something clearly in order for us to be certain that it is true. This would not be sufficient if we did not know that God exists and is not a deceiver (Fourth Replies, AT VII.171).

Charity demands that we at least try to find ways of reading Descartes's work which does not attribute to him an obvious logical error. Descartes's denial of guilt is evidence that we may be poorly interpreting Descartes when we allege that there is a problematic circle.

---

[1] Compare this formulation of the problem to Feldman 1974: 159.

[2] Citations are to the Adam and Tannery; translations are from the CSM/CSMK volumes.

Furthermore, important aspects of Descartes's project are skipped if we stop reading after the Second Meditation. The two-faculty theory of judgment in the Fourth Meditation continues to inform debates in the philosophy of mind.[3] The Sixth Meditation's argument for dualism continues to have force, at least as an argument for conceptual dualism. Descartes's idiosyncratic arguments for God's freedom to choose the eternal truths form a background for contemporary dialetheic logics. The rest of the *Meditations* is worth engaging.

Standard attempts to defend Descartes's reasoning deny that *The Meditations* involves a circle. Gewirth's seminal papers defending this claim by were followed by work by Harry Frankfurt, Anthony Kenny, Fred Feldman, James van Cleve, Janet Broughton, Jonathan Bennett, and, most recently, Michael Della Rocca. All of these interpreters take Descartes's response to Arnauld, that his reasoning is not circular, at his word.[4]

I agree with those who argue that Descartes's reasoning does not suffer from a simple logical flaw. Further, the arguments that have been adduced in Descartes's favor have been illuminating, many of them subtle and useful. Unfortunately, I believe that none of them succeed in eliminating the circle.

In this paper, I argue that Descartes does reason in a circle, but that Descartes's circularity is not problematic. My argument relies on the claim that the legitimacy of circular reasoning depends on the particulars of the case. Circular reasoning is not necessarily fallacious. There are virtuous and vicious circles.

In arguing that Descartes's reasoning is circular, but not viciously circular, I do not intend to vindicate Descartes's project. I do not accept many of Descartes's arguments. In particular, I believe that the causal argument for the existence of God is irreparably faulty, but I will not defend that conclusion, here. My goal is merely to deny that the problem with those arguments is the circularity. Critics who argue that Descartes's reasoning involves a vicious circle avoid locating the problem with his project where it belongs, on the argument for the existence of God. The real problem is with the content of the argument, not with its form; Descartes's arguments are unsound, not invalid.

To motivate my defense of Descartes's reasoning, imagine that there really were a divine guarantee. That is, imagine that there is a God who created the universe and who preserves it, with an act like that of creation, at every moment. Further, imagine that the goodness of God entails that we are never mistaken about claims that we perceive clearly and distinctly (whatever that means, and let us further suppose that we know what it means). In other words, let us grant the truth of Descartes's premises in order to examine the structure of his argument. In such a case, I believe that Arnauld's criticism would not be compelling. And, if not, then there is no structural problem with the argument.

This is not a case of an invalid form of argument appearing uncontroversial when instantiated with true premises and a true conclusion. Indeed, circular reasoning is deductively valid. The question is whether such an argument is inductively valid, and my claim is that it depends on the case.

Thus, my argument is that there are two types of responses to the problem of circularity:

1. Abandon the project;
2. Show that Descartes's reasoning is somehow not problematic
    - 2a. Show that Descartes's reasoning is not circular.

---

[3] See Rosenthal 1990.

[4] For example: "I do not really believe that Descartes knew for certain everything he said he did, e.g. his causal maxims. But this does not detract from the soundness of the general plan I am attributing to him" (van Cleve 1979: 258).

2b. Accept that Descartes's reasoning is circular, but show that the circularity is not problematic.

By arguing for a resolution of type 2b, I believe that am seeking a new position on the problem of Cartesian circularity.[5] There may be arguments in the literature that argue for solutions of type 2b, but I am not familiar with any. Further, I do not intend to argue that my position was held by Descartes. When he claims not to be guilty of circular reasoning, he may mean that he is reasoning in a circle, but that there is no need to feel guilty about it. Or, he may mean that there is no circular reasoning in *The Meditations*.

II. The analytic and the synthetic

The literature on the Cartesian circle over the last seventy years has proceeded on the premises that circular reasoning is obviously fallacious. Thus, Descartes's reasoning, if sound, must not be circular. Frankfurt notes that the problem is interpretative. "If the question were merely whether a certain piece of Descartes's reasoning is or is not circular, philosophers would hardly have found it so difficult to arrive at a generally accepted answer. Their disagreement over the validity of Descartes's argument is, however, due to their failure to agree on just what that argument is" (Frankfurt 1962: 504).

Those who deny circularity often rely, either explicitly or implicitly, on the difference between the analytic presentation in the *Meditations* and the synthetic presentation of much of the same content, which Descartes offers in the Second Replies. Euclid's *Elements* is arranged synthetically, as is Spinoza's *Ethics*. The synthetic method is formal and deductive.

Descartes devised the analytic method as part of his new philosophy. It is evident in the *Discourse*, in the *Regulae*, and, some argue, even in the *Principles*.[6] Descartes contrasts the analytic presentation, which tracks discovery, with the synthetic presentation, which tracks justification.

> Analysis shows the true way by means of which the thing in question was discovered methodically...so that if the reader is willing to follow it and give sufficient attention to all points, he will make the thing his own and understand it just as perfectly as if he had discovered it for himself... Synthesis, by contrast, demonstrates the conclusion clearly and employs a long series of definitions, postulates, axioms, theorems, and problems, so that if anyone denies one of the conclusions it can be shown at once that it is contained in what has gone before... This method is not as satisfying as the method of analysis, nor does it engage the minds of those who are eager to learn, since it does not show how the thing in question was discovered (Second Replies, AT VII.155-6)

Descartes's presentation in the *Meditations*, employing analysis, may thus not reflect the fundamental structure of his argument. Indeed, on the analytic interpretation, what Descartes provides

---

[5] But, see Frankfurt 1965 and Della Rocca 2005: 3, both of which raise, but do not pursue, questions about the nature of the circularity, and whether the standard explication of the circle even imputes circular reasoning to Descartes. Here's Frankfurt: "It is evident that Descartes is not guilty of circularity in the sense of offering an argument whose conclusion appears among its premisses. If his reasoning is circular at all, the circularity is, I take it, of a less formal variety" (Frankfurt 1965: 209).

[6] See Garber and Cohen 1982.

are not really arguments at all, but demonstrations of human psychology and the means of discovery.

> You say...that there is a vicious circle in proving effects from a cause, and then proving the cause by the same effects. I agree: but I do not agree that it is circular to *explain* effects by a cause, and then *prove* the cause by the effects; because there is a big difference between *proving* and *explaining*. I should add that the word 'demonstrate' can be used to signify either... (Letter to Morin 13 July 1638, AT II.197-8)

It hardly needs mentioning that the subtitle of the *Meditations* is "in which are *demonstrated* the existence of God and the distinction between the human soul and the body" (my emphasis).

Descartes asserts that he employed the method of analysis as a way of avoiding fundamental problems with the method of synthesis. Despite its ability to compel assent, deduction can never further one's learning.

> On the basis of their method, dialecticians are unable to formulate a syllogism with a true conclusion unless they are already in possession of the substance of the conclusion, i.e. unless they have previous knowledge of the very truth deduced in the syllogism. It is obvious therefore that they themselves can learn nothing new from such forms of reasoning, and hence that ordinary dialectic is of no use whatever to those who wish to investigate the truth of things. Its sole advantage is that it sometimes enables us to explain to others arguments which are already known (*Rules*, Rule Ten, AT X.406).

The solution to the problem of the Cartesian circle which suggests itself on reflection on these passages is that since Descartes is not providing a deductive proof of the conclusions in Meditations Four through Six, we need not worry about any circular inter-relations among the claims. Since the method of analysis is used for discovery and explanation, to augment understanding, such inter-relations are elucidating rather than fallacious.

Solutions based on taking seriously Descartes's discussion of the analytic method are compelling. They solve the problem of the Cartesian circle while being independently motivated. Indeed, my solution to the problem has strong affinities to such interpretations, as I will show.

But notice what we cede by giving up hope for a proof in the *Meditations*. We can no longer see Descartes as presenting a method for securing truth on the basis of clear and distinct perception in response to the most severe doubts. On the analytic interpretation, Descartes is merely explaining his conclusions, rather than deriving them. We are forced to think of the *Meditations* as a description of how one discovers Descartes's philosophy, rather than an argument for its truth. The project retains a psychological role, but abandons its normative epistemological role.

> All six Meditations can be seen as essentially about my nature as a thinking thing. Meditation I exhibits my mental nature by engaging in doubting and other highly focused mental acts. Meditation II describes that mental nature, and the Third Meditation establishes its reliability with respect to reaching the truth. The importance of the will in the first three Meditations culminates in the two-faculty theory of my mental nature in Meditation IV, which explains how my mental nature operates. In the Fifth Meditation I discover my innate understanding of various true and immutable natures, and in the Sixth that my mental nature is noncorporeal (Rosenthal 1990: 35).

The radical shift in our understanding of the *Meditations* caused by taking the analytic method seriously has not been fully appreciated even by those who appeal to Descartes's distinction. Thus, most defenders of appeals to the methodological distinction do not fully embrace a psychological interpretation of the *Meditations*. Instead, they look for ways to invoke that distinction while retaining the standard, normative view of the *Meditations*. As I show in the next two sections, attempts to avoid attributing circularity to Descartes by appealing to his analytic method while retaining the claim that Descartes has presented an argument or proof of his conclusions in the *Meditations*, meet serious challenges.

In what follows, I will contrast a standard reading of the *Meditations* with an analytic interpretation. By 'standard', I mean, roughly, a reading on which Descartes begins by presenting the doubts as premises in an argument whose conclusions include both the existence and goodness of God, and the veridicality of all clear and distinct perceptions, secured by the goodness of God. On the standard reading, Descartes is providing an extended, deductive argument, and the doubts function as premises requiring a further argument for their elimination. In contrast, I use 'analytic interpretation' to refer to readings on which the doubts are psychological or methodological, but on which they do not function as premises in Descartes's arguments for the existence of God or the veridicality of clear and distinct perceptions. On the standard interpretation, the goal of the *Meditations* is epistemically normative: the arguments provide reasons to take the conclusions as true. On the analytic interpretation, the goal of the *Meditations* is elucidative.

The most serious consequence of taking the analytic view seriously is that the doubts are no longer taken as providing a challenge for the aspiring knower to overcome. The role of doubts in proof would be to undermine premises. The defender of the analytic interpretation argues that since Descartes is providing a demonstration, rather than a proof, he need not show how to remove the doubts, to prove that there is no deceiver. The role of the doubts is merely as a demonstration of what it is possible to believe.

> In Meditation I Descartes's aim in doubting is simply to see whether it is possible to doubt particular sorts of things (Rosenthal 1990: 32).

Some interpreters of Descartes seek aspects of both the standard and the analytic reading. They want the standard normative aspects of Descartes's project, to see Descartes as arguing for the truth of particular claims, while understanding the *Meditations* analytically in order to avoid attributing circularity to the argument. Capturing the normative aspect of Descartes's work while maintaining the analytic reading, would be optimal; such an interpretation would fit best with what Descartes says about his work. But, if the doubts are going to have to function as premises in an argument, as hypotheses to be refuted, then the circle is going to be unavoidable. I believe that the doubts are working as premises and thus that attempts to find a way to avoid attributing circularity fail. Embracing the circle is a better option.

III. Normative conclusions from psychological premises: Gewirth and Feldman

Gewirth, in his insightful and influential "The Cartesian Circle", begins his attempt to avoid attributing circular reasoning to Descartes by distinguishing Descartes's method from his metaphysics.

> The method deals with the mind's operations upon ideas themselves, *qua* meanings directly apprehended; the norm of these operations, and of the ideas as their direct objects, is clearness and distinctness. The metaphysics, on the other hand, is concerned with the relation of ideas to the objects purportedly represented by them; its norm is truth (Gewirth 1941: 370).

In what Gewirth calls the methodological moment, clear and distinct ideas compel our assent. We can not help, as a psychological matter, but accept the content of a clear and distinct perception, no matter what its truth value. For example, when I am presented with the claim that 2+3=5, or that I exist, my assent to the claim is automatic.

> Even if I had not demonstrated [that everything of which I am clearly aware is true], the nature of my mind is such that I cannot but assent to these things, at least so long as I clearly perceive them (Meditation V: AT VII.65).

> Some of [my] perceptions are so transparently clear and at the same time so simple that we cannot ever think of them without believing them to be true (Second Replies: AT VII.145)

The immediacy of one's assent to clear and distinct perceptions can be further explained in logical terms of consistency, coherence, and necessary connection. But while clear and distinct perception is a psychological or logical matter, there is a separate question about whether our beliefs are true. The possibility of a deceiver calls into question whether we can be sure that clear and distinct perceptions are true. We can determine whether a perception is clear and distinct without ever confronting the question of whether its content is true; the two questions are independent.

> Our mind is of such a nature that it cannot help assenting to what it clearly understands... perhaps our nature is such that we go wrong even in the most evident matters (Letter to Regius, 24 May 1640: AT III.64-5).

Indeed, Descartes raises the three doubts precisely to question the claim that whatever we perceive clearly and distinctly must be true. Methodological certainty, an inability to doubt a particular claim, does not entail metaphysical certainty, as Frankfurt notes.

> Descartes's metaphysical doubt is precisely a doubt whether being false is compatible with being indubitable. His position is that as long as the demon remains a possibility, we must acknowledge that what we intuit may be false. But he also holds that we cannot doubt the truth of what we intuit while we are perceiving it clearly and distinctly (Frankfurt 1965: 212).

As Gewirth observes, statements can be true without being clear and distinct, for Descartes, and some statements can be clear and distinct without being true. In the former case, we are finite and can not comprehend some complex truths. In the latter case, God is free to determine the eternal truths.[7]

> As for the eternal truths, I say once more that they are true or possible only because God knows them as true or possible. They are not know as true by God in any way which would imply that they are true independently of him (Letter to Mersenne, 6 May 1630: AT III.24)

Descartes's idiosyncratic view on the eternal truths aside, it seems eminently plausible that we will find some claims to compel our assent without fail, even if they turn out to be false. Consider our apprehension of space as Euclidean. As psychological matter, our belief in a flat space-time is nearly impossible to doubt. When we learn about general relativity and the curvature of space, we abandon our belief in Euclidean space, as a theory of actual space, independently of how space seems to us. In cases

[7] See Gewirth 1941: 372-3; especially fn 21.

of optical illusions, the world seems to be a way in which it is not; the illusion can persist even when we know that we are perceiving an illusion. George Bealer makes a similar point regarding the axiom of comprehension, the claim that every property determines a set.

> I have an intuition - it still *seems* to me - that the naive comprehension axiom of set theory is true; this is so despite the fact that I do not believe that it is true (because I know of the set-theoretical paradoxes) (Bealer 1998: 208).[8]

Thus whether we accept Descartes's claim that God could can choose the eternal truths, or we independently accept the claim that what we perceive clearly could turn out to be false, the deceiver doubt has a force that must be overcome if Descartes is to secure the claim that our clear and distinct perceptions are true, and if he is going to use the clarity and distinctness of our perceptions to support the causal argument for God's existence.

Gewirth's attempt to save Descartes from circularity depends on rejecting the claim that Descartes's doubts impugn conclusions based on our clear and distinct perceptions. His solution is to see the *Meditations*, as containing two, overlapping narratives. The doubts are in the metaphysical moment, impugning the truth of our perceptions. But, the causal proof of God's existence is in the methodological moment, based on clear and distinct perception. The only way to impugn the certainty of the causal argument would be if one could have a serious doubt about those perceptions which are clear and distinct, if the results of the metaphysical moment could bleed into the claims in the methodological moment. And, Gewirth claims, the doubts Descartes presents are not serious enough to transfer, to undermine the certainty based on our clear and distinct perceptions.

Gewirth claims that his distinction between the metaphysical moment and the methodological moment, restricting the range of the doubts, helps explain the key fourth paragraph of Meditation Three, where Descartes implies that the cogito undergoes the metaphysical doubt, since everything does.

> But what about when I was considering something very simple and straightforward in arithmetic or geometry, for example that two and three added together make five, and so on? Did I not see at least these things clearly enough to affirm their truth? Indeed, the only reason for my later judgment that they were open to doubt was that it occurred to me that perhaps some God could have given me a nature such that I was deceived even in matters which seemed most evident. And whenever my preconceived belief in the supreme power of God comes to mind, I cannot but admit that it would be easy for him, if he so desired, to bring it about that I go wrong even in those matters which I think I see utterly clearly with my mind's eye. Yet when I turn to the things themselves which I think I perceive very clearly, I am so convinced by them that I spontaneously declare: let whoever can do so deceive me, *he will never bring it about that I am nothing, so long as I continue to think I am something*; or make it true at some future time that I have never existed, since it is now true that I exist; or bring it about that two and three added together are more or less than five, or anything of this kind in which I see a manifest contradiction. And since I have no cause to think that there is a deceiving God, and I do not yet even know for sure whether this is a God at all, any reason for doubt which depends simply on the supposition is a very slight and, so to speak, metaphysical one. But in order to remove even this slight reason for doubt, as soon as the opportunity arises I must examine whether there is a God, and, if there is, whether he can be a deceiver. For if I do not know this, it seems that I can

---

[8] Similar cases arise even for sense perception: the Müller-Lyer illusion is still effective, even when I know I am misled.

> never be quite certain about anything else (Meditation III, AT VII.35-6, emphasis added).

Gewirth explains the tension in this passage by Descartes's vacillation between the metaphysical moment and the methodological moment.

> In the first place, since only psychological, not metaphysical, certainty is possible before God's existence is known, the perceptions used to demonstrate that existence will be only psychologically certain. But, in the second place, the perception of God's existence will itself be not only psychologically but also metaphysically certain. And consequently, in the third place, Descartes's argument is not circular, for while it is by the psychological certainty of clear and distinct perceptions that God's existence is proved, what God guarantees is the metaphysical certainty of such perceptions (Gewirth 1941: 386).

The problem with Gewirth's argument is that he has not given us a reason to believe that perception of God's existence is both psychologically and metaphysically certain. He alleges that the metaphysical doubts do not apply to the premises of the causal argument, but it is difficult to see why they would not. Remember, it is still an open question whether clear and distinct perceptions track truth.

Gewirth claims that the doubts do not affect the premises in the causal argument because they are restricted to the metaphysical moment, whose norm is truth. That claim is not quite right. Using both the psychological facts about our impelled assent and the logical basis for that assent, we can distinguish among the first two doubts, on the one hand, and the third. The first two doubts (illusion and dreaming) are psychologically serious, in that one can imagine that one is mistaken about the sense properties of physical objects or about whether one is awake or asleep. We can conceive of what the world would be like, in such cases. We can adopt the doubt from the inside. Indeed, the first two doubts have a methodological role; Descartes uses them to dissuade the reader from reliance on her senses.

In contrast, the deceiver doubt can not be taken seriously, from a psychological point of view. We can not imagine what the world would be like if the eternal truths were false: if contradictions were true or two and three were not five. We can only speculate that such deception is possible. Unable to think about such a world from the inside, we have to accept the force of that doubt from the outside.

The psychological differences among the doubts helps explain a puzzling passage from the end of Meditation Six. While the deceiver doubt calls into question a wider class of beliefs, Descartes calls the dream doubt his chief one.

> The hyperbolic doubts of the last few days ought to be rejected as ludicrous. The goes especially for *the chief reason for doubting, which dealt with my failure to distinguish being asleep from being awake* (AT VII.89, emphasis added).

The deceiver doubt, while effective at calling into question the truth of those claims which are clear and distinct, is at the same time psychologically inert. It is not a real reason for doubting, from a methodological point of view, though it is clearly a reason for doubting the truth of one's beliefs. The psychological inertness of the deceiver doubt also explains Descartes's difficulty holding the doubts, at the end of Meditation One and the beginning of Meditation Two.

> This is an arduous undertaking, and a kind of laziness brings me back to normal life. I am like a prisoner who is enjoying an imaginary freedom while asleep; as he begins to suspect that he is asleep, he dreads being woken up, and goes along with the pleasant illusion as long as he can In the same way, I happily slide back into my old opinions and dread being shaken out of them, for fear that my peaceful sleep may be followed by hard labour when I wake, and that I shall have to

> toil not in the light, but amid the inextricable darkness of the problems I have now raised... It feels as if I have fallen unexpectedly into a deep whirlpool which tumbles me around so that I can neither stand on the bottom nor swim up to the top (AT VII.23-24).

The psychological differences among the doubts also explains the presence of the first two doubts in the *Meditations*. Since the deceiver can not support psychological doubt, it does not call into question the clarity or distinctness of my perceptions. Gewirth is correct that the deceiver doubt is only metaphysical, not methodological. But, in order for the reader to take the doubts seriously, Descartes presents the metaphysically weaker, but psychologically stronger doubts. The doubts infect both the methodological moment and the metaphysical moment.

So I think that Gewirth is not correct when he claims that the doubts are restricted to the metaphysical moment. Still, he is correct that the deceiver doubt, being psychologically inert, is ineffective on the methodological moment. But that just goes to show that, granting Descartes's claim that the premises of the causal argument are in fact clear and distinct, we can not doubt, as a psychological matter, those premises. It does nothing to show that they are true.

Our worries about the truth of the premises, Gewirth is arguing, arise only on the metaphysical moment. A standard reading of the *Meditations* takes the doubts to infect the premises of the causal argument. Gewirth needs a way to show that the doubts are cut off from the methodological moment, from impugning clear and distinct ideas.

His proposal is that the independence of the two moments entails that we need not merely idle or thin reasons for doubting clear and distinct perceptions, but serious ones. This claim is supported by the appeal to the analytic method. On the analytic interpretation, the doubts are not premises in an argument; the argument for the existence of God need not defeat the doubts.

Gewirth's isolation of the methodological moment from the metaphysical moment, based on the analytic method of presentation, would be more effective if the deceiver doubt did not infect the synthetic presentation. Indeed, the deceiver doubt is only implicitly present, there. The first and second postulates concern the veracity of sense perception, and do not broach the question of the existence of a deceiver there. But, in the seventh postulate, Descartes defends his reliance on clear and distinct perceptions.

> I ask [readers] to conclude that it is quite irrational to cast doubt on the clear and distinct perceptions of the pure intellect merely because of preconceived opinions based on the senses, or because of mere hypotheses which contain an element of the unknown. And as a result they will readily accept the following axioms as true and free of doubt (Second Replies, AT VII.164)

The allusion to mere hypotheses is reminiscent of Descartes's claim, in the vacillation paragraph, that the deceiver doubt is slight and metaphysical. The deceiver is lurking, even in the synthetic presentation.

Returning to the presentation in the *Meditations*, if there were valid, and well-considered reasons for doubt, they would be able to impugn the clarity and distinctness of our perceptions. In the absence of such doubts, the premises of the causal argument have not been impugned.

> Any perception which casts a metaphysical doubt upon clear and distinct perceptions must itself purport to be based upon such perceptions; the mind, conscious of the nature of its own internal rationality, can therefore indicate that if that rationality is subjected to any attack which the mind may apprehend or perceive, that attack must itself imply and exhibit that very rationality. Hence the metaphysical doubt which the mind bases upon "reasons" or rational considerations must presuppose in the very statement of it the standards of the mind's own rational processes (Gewirth 1941: 392-3).

Gewirth's argument is reminiscent of Wittgenstein's claim that doubt presupposes certainty. But it does not, as van Cleve argues.

> [Gewirth's] argument established only that clear and distinct perceptions are metaphysically certain in the sense [that] a proposition is metaphysically certain if and only if every reason for doubting it is excluded by clear and distinct perceptions... What is added when at the end of the argument we say that God's veracity and other things clearly and distinctly perceived are metaphysically certain? Just this: that we are psychologically certain not only of those propositions themselves, but also of the falsehood of every reason for doubting them. Thus, we have not advanced to a new *kind* of certainty at all. We have merely extended the class of psychological certainties (van Cleve 1979: 244).

In other words, Gewirth's argument does not allow us to derive the metaphysical certainty of God's existence; it just allows us to remove more methodological doubt. The distinction between the analytic and synthetic modes of presentation could allow Descartes to secure his conclusion, but only at the cost of abandoning the normative claim that we know that God exists on the basis of the causal argument.

The broader point is this: if the goal of the *Meditations* is, as it is standardly taken to be, to provide proofs of both the existence of God and the security of the criteria, Gewirth's solution avoids attributing circularity only at the cost of failing to meet that goal. Alternately, if the goal of the *Meditations* is, as defenders of the analytic interpretation urge, to describe the method of learning, then Gewirth's arguments that Descartes can avoid circularity by isolating the metaphysical moment from the methodological moment are otiose. Gewirth argues that the proof of God's existence is immune from the doubts, but if Descartes is not proving the existence of God, he would not need such immunity; there is no proof to secure. On the analytic interpretation, there is no circularity because there is no argument for the metaphysical certainty of Descartes's claims.

Similar solutions to the problem suffer similar fates. Feldman argues that the doubts do not actually call the premises of the argument for God's existence into question. For all we know, Descartes argues, there could be a deceiver. That epistemic, or methodological, possibility entails that my belief in the claim that two and three are five (and all other beliefs, including those which I perceive clearly and distinctly) is brought into doubt. But, the claims which are perceived clearly and distinctly are still clear and distinct; the doubts just make it the case that I don't know whether all clear and distinct claims are true. On the basis of clear and distinct reasoning, I can derive the existence of God. Still, I have no practical reason for doubting clear and distinct claims. And now, since I have proved the existence and goodness of God, the clear and distinct perceptions I used in the proof have a divine guarantee of truth. I can adopt the rule.

Just as the psychological assurance of clear and distinct perception did not, in Gewirth's argument, allow us to eliminate the possibility that our premises in the argument for God's existence are false, the metaphysical possibility that there is a deceiver still exists, even if it is no longer a practical/epistemic possibility. Feldman rejects the claim that the practical possibility has any import.

> Descartes says that reasons for doubt must be "powerful and maturely considered" (AT VII 21; cf. AT VII 459-60) and that doubt must be based upon "clear and assured reasonings" (AT VI 29). It seems unlikely that something that is practically impossible could count as a powerful and maturely considered reason for doubt, or that one could legitimately call it a clear and assured reason for doubt (Feldman 1974: 168).

Feldman says that something that is practically impossible could not be a mature reason for doubt. But, that's just wrong. Indeed, the practically impossible doubt is the only support for doubting the eternal truths. From Feldman's position, passages in which Descartes calls the eternal truths into question would be inexplicable.

Feldman relies on the premise that a claim only casts metaphysical doubt (as opposed to epistemic, or practical doubt) on another claim (as the claim that a deceiver is possible would cast on any of the premises in the argument for God's existence) if it is not the case that its negation is practically certain. On Feldman's view, all Descartes needs to do, to make the leap from the methodological level to the metaphysical level, is to show that the impossibility of a deceiver is practically certain. The psychological impossibility of doubting our clear and distinct perceptions is supposed to ensure that the deceiver doubt does not infect the premises of the argument for God's existence.

As van Cleve argues, Feldman's argument, like Gewirth's, suffers from an unjustifiable step from the methodological level to the metaphysical one.

> Feldman's conclusion is equivalent to the following: we are practically certain not only of clear and distinct perceptions, but also of the falsehood of every proposition that would cast doubts on them... Feldman's conclusion, like Gewirth's, provides no guarantee that clear and distinct perceptions are *true*. For a Cartesian, this is not enough (van Cleve 1979: 244).

For van Cleve, a 'Cartesian' demands proof that the premises of the argument for God's existence are secure from doubt. Again, the point could be made as follows: if the Cartesian wants proof, then Feldman avoids circularity only by failing to secure the metaphysical certainty of the conclusion. But, if one takes the analytic interpretation seriously, then the doubts need not be proven false; then again we have no proof.

IV. Distinguishing general and particular claims: Van Cleve and Della Rocca

The standard interpretation holds that the doubts must be overcome in order to establish the truth of the premises in the causal argument and the veridicality of clear and distinct perceptions. The problem unresolved by Gewirth and Feldman is that the deceiver doubt infects both, as seems to be the case in a natural reading of the vacillation paragraph in the Third Meditation. If Gewirth and Feldman were correct that the psychological compulsion of clear and distinct perceptions could secure Descartes's conclusions, then the doubts should not have the effectiveness that they do, especially at the end of that paragraph.

Van Cleve explains the vacillation paragraph as the result of Descartes's uncertainty about the general connection between clear and distinct perception and truth. According to van Cleve, Descartes does not doubt, for individual claims, that if they are clear and distinct, then they are true. Instead, he is expressing that he can not affirm the general connection between clear and distinct perception and truth.

Consider, with van Cleve, the claim, 'I am certain of the truth of clear and distinct perceptions'. It may be interpreted in either of two ways.

> A. For all p, if I clearly and distinctly perceive that p, then I am certain that p.
> B. I am certain that for all p, if I clearly and distinctly perceive that p, then p.

Van Cleve alleges that Descartes vacillates on the status of the propositions which are clear and distinct because he holds (and is justified in holding) A, but not B.

The appearance of inconsistency is removed if we see Descartes as being uncertain not of

> particular propositions that he clearly and distinctly perceives, but only of the general connection between clear and distinct perception and truth. What he shows us in the paragraph is that at this stage in the *Meditations*, (A) is true of him, but (B) is not (van Cleve 1979: 246).

Thus, on the one hand, Descartes is certain of the eternal truths, including the premises of the argument for the existence of God. And, on the other hand, he is not certain of the entailment from clear and distinct perceptions to the truth of the content of those perceptions, in general.

According to van Cleve, the doubts only call into question our knowledge that we are certain, not that certainty itself; they are second-order doubts. Van Cleve avoids attributing a circle to Descartes by arguing that the proof of God's existence removes only those second-order doubts. Descartes, according to van Cleve, is arguing:

> The argument for the existence of God is based on my clear and distinct perceptions.
> I have doubts whether clear and distinct perception ensures certainty, even though it does.
> The goodness of God ensures that I know that my clear and distinct perceptions, which always were secured even though I did not know it, ensure certainty.

Van Cleve's claim, in short, is that Descartes always was certain of all clear and distinct propositions, since A is true of him at the beginning of the *Meditations*. He was just uncertain that he was certain of them.

> By falling under proposition (A), (that is the C&D rule), he becomes certain of premises from which he eventually derives proposition (A) itself. But since he does not have to *use* proposition (A) at any step along the way, there is no circle... After he has proved proposition (A) Descartes can give the following reply: "Those premises are things I knew for certain. The proof of this is that I perceived them clearly and distinctly, and whatever I so perceive is certain." (Van Cleve 1979: 248).

On van Cleve's interpretation, the *Meditations* is an attempt to justify claims of which we were already certain, rather than to become certain of anything further.

It is odd that van Cleve uses 'am certain' in A and B, rather than 'know'. 'Knowledge' is a success term; from the fact that I know that p, it follows that p. In contrast, 'certainty' is ambiguous. From the fact that I am certain that 2+3=5, it may or may not follow that 2+3=5. Della Rocca distinguishes between normative certainty, which like 'knowledge' entails truth, and psychological certainty, which does not. On the psychological reading, 'p is certain' means that 'I have no doubt that p'. Our psychological certainty does not entail the truth of that of which we are certain.

There is no serious question whether Descartes accepts that clear and distinct perceptions have psychological force which makes doubt impossible. When I think that 2+3=5, I can not doubt it. But, the First Meditation discussion raises the question of whether there can be a deceiver, and concludes that such a deceiver can make propositions which appear clear and distinct false. So, on the psychological reading of 'certain', van Cleve is correct that A holds but is unhelpful in defeating the deceiver. Further, on the psychological reading, B does not hold of Descartes. But, on the standard reading, B is not Descartes's quarry. Descartes wants more than psychological certainty of the security of his clear and distinct perceptions; he wants to know that they are true.

Van Cleve's use of 'am certain' is an attempt to strengthen a similar argument from Anthony Kenny. Kenny argues that the proof of the existence and goodness of God only ensures a general principle, and not particular ones.

> There is no circle in Descartes's argument. The clear and distinct perceptions used in the proof of God's existence are perceptions of particular propositions, such as that ideas cannot be more perfect than their archetypes. The veracity of God is used to establish not any particular clear and distinct perception, but the general proposition that whatever I clearly and distinctly perceive is true (Kenny 1970: 690).

Kenny gets to his conclusion by distinguishing the following propositions.

A'. For all p, if I clearly and distinctly perceive that p, then I cannot doubt that p.
B'. I cannot doubt that for all p, if I clearly and distinctly perceive that p, then p.

But A' is too weak to support Kenny's conclusion; our inability to doubt is a psychological fact about us, not a normative claim about our knowledge. Van Cleve's strengthening of Kenny's distinction is an attempt to make the consequent of A normative. So, a proper reading of van Cleve entails reading 'am certain' with normative force. Van Cleve thus claims that not only are we psychologically certain of Descartes's argument for the existence of God, but that its truth is secured, before we defend the criterion of clear and distinct perception. Van Cleve's solution to the circle is that A" holds of Descartes at the beginning of the Third Meditation; the proof of clear and distinct perceptions only establishes B".

A". For all p, if I clearly and distinctly perceive that p, then I know that p.
B". I know that for all p, if I clearly and distinctly perceive that p, then p.

Van Cleve's solution agrees with the analytic interpretation of the *Meditations* by putting the doubts aside for A". But, van Cleve goes on to show that the doubts are infecting B", and must be defeated. If van Cleve were adopting the analytic interpretation, such claims would not be necessary; on the analytic interpretation, the role of the doubts was merely to explore the psychology of philosophical methodology, not to be put as premises to be overcome in an argument.

Della Rocca, in a recent paper, agrees with van Cleve that Descartes starts the *Meditations* by presupposing normative certainty. Della Rocca argues that the purpose of the criterion is not to secure the premises of the argument for the existence of God, which would make the argument circular. Instead, the criteria are used to achieve retrospective certainty.[9]

> Descartes allows that one can doubt a clear and distinct idea when one is no longer attending to it... Although at the moment of clearly and distinctly perceiving the idea we are, on this view, normatively certain of the truth of the idea, we can after the fact wonder whether this idea that we previously clearly and distinctly perceived is true and we can wonder whether our clear and distinct ideas generally are true. At these later times, we are not normatively certain of the claim in question, though we were normatively certain of it earlier... Once Descartes carries out the argument for God's veracity, Descartes has good reason to believe that his clear and distinct ideas are true and thus Descartes can...have normative certainty of claims that he no longer clearly and

---

[9] Willis Doney had argued that Descartes's arguments for doubt infect clear and distinct perceptions only when they are brought forth in memory (Doney 1955). Intuitions, when they are perceived, are not subject to the same doubts. Descartes does call even direct intuitions into doubt, so Doney's interpretation seems implausible. Frankfurt 1962 is widely taken to have refuted Doney's interpretation. Della Rocca's claim that the criteria of clear and distinct perceptions is used only to secure "retrospective certainty" is an heir to Doney's claim.

distinctly perceives (Della Rocca 2005: 11-12).

Della Rocca and van Cleve thus share the claim that Descartes emerges from the First Meditation with an ability to achieve knowledge or normative certainty in tact. The argument for the existence of God is thus immune from deception; all Descartes needs from the criterion is to know that the proof is immune from deception. Concomitantly, the criteria of clear and distinct perception only protect second-order knowledge, leaving first-order knowledge alone, since it was already protected. Della Rocca argues that Descartes should taken as an externalist about epistemology, having knowledge despite not, at first, knowing that he does; indeed Della Rocca only avoids attributing circularity to Descartes by claiming that the second-order knowledge he gets from the causal argument and the subsequent security of the criteria is not attributable to Descartes before he makes those derivations.[10]

Della Rocca's claims entail that the *Meditations* does not contain an argument against skepticism, since Descartes's first-order claims were never really called into question. Van Cleve and Della Rocca thus secure the normative conclusions of the *Meditations* by claiming that they were never impugned by the doubts. On their reading, Descartes is claiming only that he can not know that he knows anything until he prove that God exists. The standard normative reading of the *Meditations* is that Descartes is claiming that he can not know anything until he proves that God exists.

For van Cleve to avoid attributing circular reasoning while ascribing a normative conclusion to Descartes, clear and distinct perception must track truth. And, that is still an open question. Della Rocca claims that the justification of normative certainty has nothing to do with clear and distinct perception, which is not established until the Fourth Meditation; Descartes's normative conclusion follows just from the fact that the premises of the argument for God's existence are apparent.

Van Cleve and Della Rocca avoid attributing circularity to Descartes, but fail to account for the extreme doubt at the end of the vacillation passage.

> For if I do not know [that God exists and is not a deceiver], it seems that I can never be quite certain about anything else (Meditation Three, AT VII.36).

Descartes is not claiming to have knowledge in individual cases, but lacking knowledge of a general claim. Instead, he is wavering on the strength of the deceiver doubt. He is compelled psychologically by clarity and distinctness, and yet compelled in the other direction by the deceiver doubt. The inconsistency in the vacillation paragraph is just the same as at the strain he expresses at the beginning of the Second Meditation. We know that we must take the deceiver doubt seriously, but we can not take it seriously, as a psychological matter.

Van Cleve admits that his interpretation does not square with the last sentence of the vacillation paragraph, but says that the sentence is an embarrassment for any interpretation of Descartes.

> It must be said...that the final sentence of this paragraph...is an embarrassment for almost any interpretation of Descartes. Here he digs himself into a put so deep there can be no climbing out. Some interpreters...regard this sentence as an aberration (van Cleve 1979: 257).

---

[10] Thus, if Descartes were to hold a KK thesis, Della Rocca's argument would fail to avoid attributing circularity to Descartes. For, then Descartes's knowledge of the premises would entail that Descartes already possessed the second-order knowledge that Della Rocca claims only the subsequent arguments secure. Whether Descartes holds a KK thesis is a topic of debate. (See Rule 1 of the *Discourse*, Part 2, AT VI.18).

Della Rocca, like Gewirth, argues that the doubts of the First Meditation are not serious enough to eliminate normative certainty. The psychological certainty of clarity and distinctness is enough to establish normative conclusions.

> Descartes's doctrine of compelled assent seems...to be normative in character and thus the kind of compelled assent that currently intuited clear and distinct ideas involve may well be compelled assent by virtue of good reasons. Thus...the passages in which Descartes seems to doubt all do not - in part for reasons given explicitly by Descartes himself - support the view that, as he enters the theological argument he has no normative certainty" (Della Rocca 2005: 13).

Della Rocca argues that van Cleve's interpretation of the last sentence of the vacillation paragraph is unjustifiably despairing. First, there is the little matter of the 'else' at the end, which Haldane and Ross had omitted, leading scholars to read the sentence too strongly, historically. Second,

> This passage...can be read as elliptical for "I can never be quite certain about anything else *as long as I am not currently perceiving those things clearly and distinctly*." The passage from the Seventh Replies...shows that this kind of qualification is one that Descartes is quite willing to make. Van Cleve does not appreciate this point and so he makes unnecessarily heavy weather of this passage (Della Rocca 2005: fn 25).

> The passage to which Della Rocca refers is the following:

> [U]ntil we know that God exists, we have reason to doubt everything (i.e. everything such that we do not have a clear perception of it before our minds, as I have often explained) (Seventh Replies, AT VII.546).

But that passage does not necessarily support Della Rocca's interpretation. The vacillation paragraph shows that an inability to doubt a claim does not entail that claim's truth. On the standard reading, that is the point of the deceiver argument: to show that clear and distinct perceptions are liable to doubt unless we have reasons to think that they are secured. Doubt about certain claims is psychologically impossible, and yet there remains a worry about whether all of my clear and distinct perceptions are true.

Again, Della Rocca's claim that Descartes has normative certainty going in to the causal argument would be supported by the analytic interpretation of the *Meditations*, if all doubts were cast aside. But, since Della Rocca, like van Cleve, insists on continuing to take the doubts seriously at some level, it is clear that he continues to see them as premises in an argument. Further, Della Rocca's insistence on seeing the *Meditations* as establishing normative certainty as its goal, as the standard reading demands, rather than as an explication of the method of discovery, the psychology of philosophy, shows that he can not rely on the analytic interpretation to dismiss the doubts.

In contrast to both van Cleve and Della Rocca, I think it would be a mistake to minimize the importance of the strong arguments for universal doubt. The key sentence in the vacillation paragraph seems to be a restatement of the seriousness of the doubts, of the importance of proving the existence and goodness of God in order to eliminate those doubts. In order to deny that Descartes reasons in a circle, Della Rocca and van Cleve weaken the scope and seriousness of the doubts. If we are forced to continue to take the doubts seriously in the Third Meditation, we had better find a way to become comfortable with circular reasoning.

To see the structure of the argument in favor of circular reasoning that I am about to present, consider Della Rocca's neat analysis of the Gewirth/Feldman attempt to avoid attributing circularity to

Descartes. He starts with a general claim.

> (A) If Descartes is, prior to the conclusion of his theological argument in the Third and Fourth Meditations, at most merely psychologically (and thus not normatively) certain of propositions in general, then Descartes can not by means of argument go on to acquire normative certainty of some propositions (Della Rocca 2005: 4).

Gewirth and others accept the antecedent of A, and deny the general principle in order to argue that there is a way to avoid the circle. Della Rocca argues that these standard attempts fail to impugn A, and that there is a way to avoid circularity without denying A. I agree with Della Rocca that accepting the antecedent of A will lead one to interpret Descartes's reasoning as circular, but I want to argue that such interpretations need not be problematic.

In order to locate my argument more precisely, and clarify some of the relations among the various proposals we have already seen, I expand Della Rocca's principle.

> α. Descartes is, prior to the conclusion of his theological argument in the Third and Fourth Meditations, at most merely psychologically (and thus not normatively) certain of propositions in general.
> β. If Descartes is, prior to the conclusion of his theological argument in the Third and Fourth Meditations, at most merely psychologically certain of propositions in general, then there is no way to avoid circularity in his argument.
> γ. If there is no way to avoid circularity in Descartes's argument, then Descartes can not by means of argument go on to acquire normative certainty of some propositions.
> Therefore, Descartes can not by means of argument go on to acquire normative certainty of some propositions

Gewirth, Feldman, and others[11] accept α and γ, denying β. Della Rocca and van Cleve deny α, accepting β and γ; the conjunction of the latter two is Della Rocca's original (A). Della Rocca provides no direct argument for (A), though, arguing only that standard attempts from Gewirth and others fail to impugn its plausibility. I accept α and β, and deny γ. In fact, circular reasoning is not fallacious, generally; it depends on the circle.

In the next section, I will discuss four examples of what I regard as virtuous circles, in order to defend the circular reasoning in Descartes's work.

V. Embracing the circle

My first example of legitimate circular reasoning concerns the adoption of intensional idioms into our most austere theory. As is well-known, in "Two Dogmas of Empiricism," Quine denies that there are meanings, or Fregean senses. His larger goal is to impugn the analytic/synthetic distinction, which he takes to rely on a conception of truth in virtue of meaning; that portion of the argument is mainly beyond our interest, here.[12]

---

[11] Della Rocca also cites Doney and Curley.

[12] QD1. If there is an analytic/synthetic distinction, there must be a good explanation of synonymy.

Quine argues against the existence of meanings by showing that the Fregean theory of sense depends on a substantial notion of synonymy. If we are to posit an object, it must have clear identity conditions; we should admit no entity without identity. The identity conditions for meanings are synonymy; two terms have the same meaning if and only if they are synonymous. But in order to explain synonymy, one has to appeal to concepts which presuppose a concept of meaning or synonymy. All attempts to characterize synonymy are inter-related. Any attempt to define synonymy leads one into a circle.

Specifically, Quine tries three grounds for defining synonymy:

S1. Logic
S2. Dictionary definition
S3. Interchangeability (substitutivity) salva veritate

For S1, Quine considers Carnap's use of meaning postulates. For Carnap, if we want to say that 'Fx' and 'Gx' are synonymous predicates, within the theory, we add an axiom, or semantic rule:

$(x)(Fx \equiv Gx)$

For example, we can introduce the synonymy of 'bachelor' and 'unmarried man' by taking as an axiom of our best theory:

$(x)[Bx \equiv (Mx \bullet \sim Wx)]$

Carnap's proposal involves devising a list of two types of semantic rules. The first type specify, recursively or otherwise, which sentences are analytic. The second type says that those statements specified are among the truths. With the rules as part of our theory, we then define analyticity as 'true according to the semantic rules'. Such semantic rules, once adopted, assure us that all substitutions of synonymous expressions will maintain analyticity. That is, the synonymy rules preserve substitutivity salva analyticity within our theory, by definition. But the rules do not explain why certain semantic rules are picked out as special.

> Semantical rules determining the analytic statements of an artificial language are of interest only in so far as we already understand the notion of analyticity; they are of no help in gaining this understanding (Quine 1953: 36).

In other words, we introduce semantic rules for synonymy on the basis of a prior understanding of which terms are to be taken as synonymous, which sentences are to be taken as analytic. We thus define synonymy in terms of analyticity, in order to explain analyticity in terms of synonymy. That's a

---

QD2. The only ways to explain synonymy are by logical rules, definition, or interchangeability salva veritate.
QD3. Logical rules can not explain synonymy.
QD4. Definition can not explain synonymy.
QD5. Interchangeability can not explain synonymy.
QD6. Thus, there is no good explanation of synonymy.
QDC. Thus there is no analytic/synthetic distinction.

circle.

Saying that analytic statements are true ‘by definition’, S2, also presupposes, rather than explains, synonymy. Terms are found to be synonymous, not made so by fiat. The lexicographer is a sociologist, who reports synonymy, and so can not ground it. Explication, which adds clarifying information to a definition, relies on other, preexisting synonymies. There are a small number of exceptions: definitions by stipulation (e.g. when scientists name a planet or molecule), but these exceptions are rare. Taking synonymy as dictionary definition presupposes the community’s understanding of which terms are synonymous. But, the community takes terms as synonymous on the basis of something like dictionary definitions (or other community standards). Again, we have a circle.

The most interesting characterization of synonymy as substitutivity, S3, comes from linguistics. The idea here is that terms will be synonymous when they can be substituted for each other without changing truth values. You can substitute ‘unmarried man’ for ‘bachelor’ in ‘a bachelor is not married’, and other related expressions. There are silly examples of failure of substitutivity, which we can put aside.

> When I graduated, I received a bachelor of arts diploma.
> ‘Bachelor’ has eight letters.

The more interesting worry concerns examples like ‘creature with a heart’ and ‘creature with a kidney’, which can be exchanged in most contexts, since they are coextensive, true of the same creatures. In order to rule out such substitutions, we must strengthen the condition for substitution. A natural attempt would insist on substitutivity salva analyticity: two terms are synonymous if they can be substituted for each other while maintaining the analytic or synthetic nature of the expression in which they are substituted. But again, we are met with a circle: defining synonymy in terms of analyticity in order to define analyticity in terms of synonymy.

One last attempt would be to make the condition for substitution modal. The idea is that we can isolate synonymous expressions as those whose identities are necessary.

> Necessarily, bachelors are unmarried men.
> Necessarily, anything plain is unadorned.

Again, such sentences are circular, in that they explain one intensional idiom (synonymy/analyticity) in terms of another (modality). Quine says that this attempt has an “air of hocus-pocus.” All of the attempts S1-S3, like Descartes’s project, are circular.

> Our argument is not flatly circular, but something like it. It has the form, figuratively speaking, of a closed curve in space (Quine 1953: 30).[13]

Quine wants to explain synonymy without appealing to any other intensional contexts. He wants a reduction of the intensional. But it seems impossible to analyze meaning reductively at all.

---

[13] I attribute Quine’s hedge on calling the definitions circular, like Frankfurt’s concern about Descartes’s reasoning, to the size of the circle. “It is evident that Descartes is not guilty of circularity in the sense of offering an argument whose conclusion appears among its premisses. If his reasoning is circular at all, the circularity is, I take it, of a less formal variety. I shall not undertake to define this variety of circularity but shall assume that, in an intuitive way at least, its nature is sufficiently clear for the purposes of my essay” (Frankfurt 1965: 209).

In contrast, we can distinguish between virtuous and vicious circles. We have a set of inter-theoretically linked intensional terms: analyticity, meaning, synonymy. We can adopt the whole group by appealing to their systematic virtues for the intensional idioms themselves. Jerrold Katz takes this approach, in what he calls a non-reductive or autonomous theory, defining sense in terms of sense properties.

> (D) Sense is that aspect of the grammatical structure of sentences that is responsible for their sense properties and relations (e.g. meaningfulness, meaninglessness, ambiguity, synonymy, redundancy, and antonymy) (Katz 2004: 17).

By adopting the autonomous theory of sense, defining sense in terms of its sense properties, we give up the demand for a reductive definition, and take the whole family of related notions at once.[14] While the definition is indeed circular, the circle is not vicious. It is large enough to allow substantial characterizations of the relations among the various intensional properties.

Of course, the autonomous theory of sense is not acceptable to Quine. But Quine's objections are not structural, or formal. Quine has more substantial arguments against meanings. His real arguments against meanings are Okhamist, concerning whether they need to be introduced in order to account for behavior.

> The question of assuming intensional notions in our theory comes down to the question of whether they would play a useful role in a theory that meets the test of prediction. That is where the doubts come (Quine 1990: 198).

Similarly, the real arguments against Descartes's project, the ones underlying many of the critics who attribute circularity to his method, are not about the structure of the argument, but the content. I'll return to this point.

My second example concerns mathematical theories. Consider whether we should adopt mathematical axioms into our best theories of the world, and imagine, for the sake of argument, that we can write those scientific theories without them. Still, we might want to adopt the mathematical theories on their own merits.

The standard account of our knowledge of the various diverse theorems and theories of mathematics is that we derive all the particular claims from some basic set of axioms. Then, we take the axioms as true, or useful, or as meeting whatever epistemic norm you prefer. Once we have an account of our knowledge of the axioms, an account of our knowledge of the theorems is no more problematic; we just derive them using plain old logic.

The problem with the standard account, and where the circle arises, is that we choose the axioms on the basis of the theorems. No less an astute logician than Bertrand Russell noticed this phenomenon.

> When pure mathematics is organized as a deductive system - i.e. as the set of all those propositions that can be deduced from an assigned set of premises - it becomes obvious that, if

---

[14] The idea of a adopting a family of related notions is derived, in fact, from Quine's adoption of the family of notions of logical truth: equivalence, implication, consistency, etc. See Quine 1986: 48-49. We could take this family of related logical notions as a fifth example of a virtuous circle. For a sixth, we could look to David Chalmers' defense of property dualism in response to the hard problem of consciousness, and the adoption of the language of conscious mental states in the absence of physicalistic reductions.

> we are to believe in the truth of pure mathematics, it cannot be solely because we believe in the truth of the set of premises. Some of the premises are much less obvious that some of their consequences and are believed chiefly because of their consequences. This will be found to be always the case when a science is arranged as a deductive system. It is not the logically simplest propositions of the system that are the most obvious, or that provide the chief part of our reasons for believing in the system (Russell 1924: 325).

Our adoption of mathematical theorems is not based on insight into fundamental axioms. We just take the whole system at once. Our justification for adopting mathematical axioms depends on what theorems they yield. But our justification of adopting the theorems depends on the axioms. Again, the reasoning is circular. But, it is a virtuous circle.

Russell's insight into justification in mathematics was extended, with explicit mention of virtuous and vicious circles, by Nelson Goodman (my third example) and then adopted, as a methodology, and named, reflective equilibrium, by John Rawls (my fourth example). Goodman noticed similarities between our justifications of induction and deduction.

> How do we justify a *de*duction? Plainly by showing that it conforms to the general rules of deductive inference. An argument that so conforms is justified or valid, even if its conclusion happens to be false... Principles of deductive inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences we actually make and sanction. If a rule yields unacceptable inferences, we drop it as invalid. Justification of general rules thus derives from judgments rejecting or accepting particular deductive inferences (Goodman: 1979: 63-4).

According to Goodman, we do not have a priori insight into the correctness of some abstract, general principles of deduction. Instead, we have simple beliefs about which inferences are acceptable. We formulate deductive principles which accord with these inferences. We accept inferences which follow the deductive principles we construct. We justify the particular inferences by the general deductive principles, and the deductive principles by their individual instances. That's another virtuous circle.

Similar remarks hold for induction. Consider the introduction of a the term 'tree'. We see that there are some similarities in our environment (elms, maples, oaks). We introduce a general term, 'tree' to apply broadly to elms and maples and oaks, and not to apply to mountains or cats or grass. Once we introduce this term, we look for some explanation of what makes something a tree, we look to determine some essence or unifying principles. Once we have found unifying principles, we can use them to determine whether borderline cases (e.g. pomegranate shrubs, azaleas, geraniums) are, in fact, trees. In some cases, we will discover that terms we have chosen do not apply to all the objects we thought they did. So, 'fish' does not apply to whales, even if we originally introduced it to apply to all sea creatures. Scientists discovered regularities and uniformities among more hidden properties of mammals and other fish which override their more obvious properties.

> An inductive inference, too, is justified by conformity to general rules, and a general rule by conformity to accepted inductive inferences. Predictions are justified if they conform to valid canons of induction; and the canons are valid if they accurately codify accepted inductive practice (Goodman: 1979: 64).

Goodman's account of both deductive and inductive justification is clearly and unapologetically circular. We justify our particular inductive or deductive claims or beliefs in terms of general principles from which they follow. We justify our general principles in terms of the specific claims they yield.

> This looks flagrantly circular... But this circle is a virtuous circle. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend*. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either (Goodman: 1979: 64).

The virtuous circle that Goodman defends has come to be known, in the work of John Rawls and others, as reflective equilibrium. In *A Theory of Justice*, Rawls develops a normative ethical theory, beginning without any presumption of either first principles of justice or stable particular claims. Instead, he begins with tentative principles and raw intuitions and works up to theories of justice and considered judgments. These are adopted together, as mathematical theorems and their axioms, intensional idioms, and logical principles.

Once we admit that there are virtuous and vicious circles, we are compelled to characterize this difference, present some method of distinguishing between the two. I think that there will be no clear line of demarcation. But, the acceptability of circular characterizations will depend on the elements involved. Determining which circles are acceptable and which are not s a corollary of the demarcation problem in the philosophy of science. I shall not solve that problem here, but I can give an example. Here is a list, taken from Burgess and Rosen's *A Subject with No Object*.

> BR1. Correctness and accuracy of observable prediction;
> BR2. Precision of those predictions and breadth of the range of phenomena for which such predictions are forthcoming, or more generally, of interesting questions for which answers are forthcoming;
> BR3. Internal rigour and consistency or coherence;
> BR4. Minimality or economy of assumptions in various respects;
> BR5. Consistency or coherence with familiar, established theories;
> BR6. Perspicuity of the basic notions and assumptions;
> BR7. Fruitfulness, or capacity for being extended to answer new questions. (Burgess and Rosen 1997: 209).

The challenge for the Cartesian, on my view, would be to decide whether the adoption of both the criteria of clear and distinct perception and the argument for the existence of God fulfill enough of the criteria BR1-7 to satisfy us that they should be accepted. I would argue that they do not; my guess is that Descartes would argue that they do. Those arguments, though, do not rest on whether there is some kind of logical fallacy, or error in reasoning. They rest on the particular arguments, for example the causal argument for the existence of God.

VI. Coherentism and the analytic mode

One worry about embracing the circle is that it seems to commit Descartes to epistemic coherentism. The standard reading of the *Meditations* is foundationalist. Epistemic foundationalism is the claim that we accept some basic claims as secured, and from them, using some sorts of acceptable transmission rules, governed presumably by a demand for consistency, we derive the rest of our justified beliefs. Coherentism, in contrast, is committed only to the transmission rules. Foundationalist theories have axioms and rules of inference; everyone needs rules of inference.

My interpretation of Descartes avoids ascribing coherentism to his work for three related reasons. First, the fact that claims are in part justified by each other does not impugn the immediacy of our apprehension of their truth. Descartes is committed to self-evident principles: the cogito, the premises in the theological argument, the eternal truths. We still have intuitions about the nature of the world, and especially the eternal truths.

Second, the premises of the theological argument and the criterion are both taken as immediately justified, even if the doubts later impugn our beliefs. The arguments against the doubt, for the security of the criteria, rehabilitate what was taken as justified in the first place; they merely deflect a challenge. In contrast, consider the more subtle question of tuning axioms of set theory to particular theorems, or Rawls's considered judgments to his principles of justice. In those cases, attributions of coherentism are applicable, since neither the particular claims nor the general principles are set from the start.

Third, the allegation of circular reasoning only really applies to the analytic mode of presentation in the *Meditations*, not to the synthetic mode in the Second Replies. In the synthetic mode, the doubts function to remove reliance on sense information, not to call *everything* into question. To claim that Descartes's work is coherentist is to ignore the clearly foundationalist versions of the same content.

Still, it may be the case that the *Meditations* are, because of the analytic mode of presentation, liable to attributions of some kind of coherentism. The deductive structure of synthetic presentations serves the foundationalist well, isolating axioms as first principles, and using unassailable logical inference as a method of deriving all particular claims. By rejecting the synthetic mode of presentation, Descartes dismisses its deductive nature.

I have taken a strong view of the analytic interpretation of the *Meditations* on which the doubts do not function at all in premises for an argument, and the goal of the *Meditations* is not a deductive argument for particular conclusions. The defender of the analytic mode is likely to want to find a middle ground between just denying that the doubts of the First Meditation infect the premises of the causal argument and arguing that there are no normative claims to be found in the *Meditations* at all. Such an interpretation of the analytic mode of presentation would likely have much in common with my claim that Descartes's reasoning is virtuously circular. Indeed, my interpretation and the defender of the moderate analytic interpretation agree that there are normative claims to be found in the *Meditations*, and that the reasoning toward those claims is not fallacious. We would disagree about whether the First Meditation's doubts are to be taken literally into the Third and Fourth Meditation, as premises to be defeated. I believe that they are. The defender of the moderate analytic interpretation may not.

Still, the defender of the moderate analytic interpretation must find some way to deny that the doubts are in effect, while maintaining that the arguments in the Third and Fourth Meditation have normative force. It seems unlikely to me that any argument for the normative force of the conclusion of the causal argument would not transfer back to the original doubts. That is the lesson to be learned from the work of Gewirth, van Cleve, and Della Rocca.

VII. Conclusions

My claim that Descartes can embrace the circle is not, at root, exegetical. I am not claiming that Descartes in fact accepted circular reasoning. Indeed, his response to Arnauld indicates that he thought that circular reasoning was illicit. I suspect that Descartes had a narrow notion of circular reasoning, one on which any circularity would indeed be illicit. But, Descartes introduces the *Meditations* with a hint that circular reasoning is not automatically fallacious.

> I have always thought that two issues - namely, God and the soul, are chief among those that ought to be demonstrated with the aid of philosophy rather than theology. For although it suffices for us believes to believe by faith that the human soul does not die with the body, and

> that God exists, certainly no unbelievers seem capable of being persuaded of any religion or even of almost any moral virtue, until these two are first proven to them by natural reason... Granted, it is altogether true that we must believe in God's existence because it is taught in the Holy Scriptures, and, conversely, that we must believe the Holy Scriptures because they have come from God. This is because, of course, since faith is a gift from God, the very same one who gives the grace that is necessary for believing the rest can also give the grace to believe that he exists. Nonetheless, this reasoning cannot be proposed to unbelievers because they would judge it to be circular (Letter of Dedication, AW VII. 35).

More exegetical work would need to be done to determine the extent to which Descartes thought that circular reasoning is acceptable. My claim in this paper is that embracing the circle would give Descartes a way to defend his claim that his project is not committed to illicit reasoning, while at the same time integrating the doubts, the argument for the existence of God, and the divine guarantee in the way that the *Meditations* seems to do, and in a way that the deniers of a circle have difficulty explaining.

For Descartes, intuitions, protected by the divine guarantee, are taken as infallible. Embracing the circle, we may lose the absolute, normative certainty that Descartes thought we could have. The question, from the perspective of embracing the circle, is whether to take on the whole project of taking one's clear and distinct beliefs to be justified. Descartes turns to God to guarantee their security.

Alternatively, we might argue to the acceptability of clear and distinct perceptions from the repellence of skepticism: If we can not accept the perceptions we take as most clear and distinct, then we have no route to knowledge. But, we obviously know something. So, we should accept our most clear and distinct perceptions. Such an approach to justifying one's clear and distinct perceptions is far from Descartes's project, but it can provide us a way of understanding Descartes's work without the theological elements.

Della Rocca takes an understanding of how to do epistemology without God as a goal for his interpretation of the *Meditations*. Like the approach of the previous paragraph, I don't take Della Rocca to be presenting an exegetically serious conclusion, either. These are both ways of salvaging significant portions of Descartes's work while abandoning its theological aspects. We can accept those elements of the system that are central to the further purposes of seeking firm and lasting knowledge: the psychological force of clear and distinct perception, and its reliability.[15] We can abandon the invocation of God as the securer of the method of clear and distinct ideas; it its place we just put a dogmatic rejection of the skeptic. The method that we can glean from the *Meditations* is to reflect, stepwise, to make lists. As long as we have some self-evident propositions, then we are not committed to coherentism, and we can salvage a fallibilist foundationalism, adopting Descartes's insights into the relation between clear and distinct ideas and truth.[16]

Nothing I have said about embracing the circle vindicates the sort of reasoning found in the letter of dedication. Such a circle is too closed, the content too contentious. To evaluate such beliefs, we would have to see whether they are reliable; the evidence is weak. But, if we believed that God played a substantial explanatory role, if we believed that the God hypothesis were perspicuous, parsimonious, and predictive, we might believe differently.

---

[15] There is no space here to argue for the reliability of intuition, but see Bealer and Sosa.

[16] This is not a defense of a Carnapian internal/external perspective, in which we are making pragmatic decisions whether to adopt a language. Our theology-free Descartes is making serious metaphysical claims.

Observations on Cooperative-Learning Group Assignments

Published in *The APA Newsletter on Teaching Philosophy* 9.2, Spring 2010: 2-5.

Russell Marcus

Chauncey Truax Post-Doctoral Fellow

Department of Philosophy, Hamilton College

198 College Hill Road

Clinton NY 13323

rmarcus1@hamilton.edu

(315) 859-4056 (office)

(315) 381-3125 (home)

(315) 335-4711 (mobile)

October 14, 2009

Abstract:

Good cooperative-learning lessons are naturally appropriate in philosophy classes, and can be productive and fun. Careful attention to the formation of groups can facilitate good cooperative lessons. In this article, I distinguish methods of dividing a class into groups from the resultant distributions of students in those groups. I defend uses of random assignments of students to groups.

Introduction

Cooperative learning, known more colloquially as group work, plays an uneasy role in the undergraduate classroom.[1] It can seem inefficient to the instructor who is pressed to complete a syllabus. It can seem insufficiently sophisticated to students who are reminded of elementary-school lessons. And, it can be seen as an excuse for not preparing a proper lecture. Yet, students in all disciplines, at all levels, benefit from learning to present their work orally, and to their peers. Cooperative-learning exercises require not only discussing one's own ideas, but listening to those of others. They demand active engagement in class.

Both presenting one's ideas and listening to those of others are essential to philosophical practice. What other academic discipline, outside of the arts, has so much of its canon written in dialogue? While there is very little written about how to teach cooperative-learning philosophy classes, many philosophy instructors integrate such lessons into their courses. Those who use group work quickly discover that it is hard work to design a successful cooperative lesson. Such lessons require significantly more preparation than ordinary lectures. I use cooperative-learning exercises in a range of undergraduate philosophy classes, including Logic as well as classes in which discussion comes more naturally.[2]

Well-constructed cooperative-learning exercises may be distinguished from simple group work by attention to four factors: 1. Careful distribution of students into groups; 2. Assignments of specific roles and responsibilities to each member of the group; 3. Specific and attainable objectives; and 4. A balance of emphasis on both group dynamic and individual accountability. This article explores the first factor. Dividing students into groups is an area in which instructors are liable to make the most deleterious errors. Careful consideration of some subtle factors can facilitate successful lessons. In dividing a class into groups, there are two main factors to consider: the method by which students are assigned to groups and the range of ability levels in each resultant group. I discuss both, starting with the latter. In the end, I defend uses of random assignments of students to groups.

Group Types

All groups of students will be heterogeneous according to many factors, some of which are likely to be irrelevant to any philosophy lesson (e.g. students' heights), others of which may be salient (e.g. their reading abilities). Typical research on cooperative learning tends to focus on grouping according to characteristics like gender and race. Indeed, many cooperative-learning techniques were developed in order to facilitate interactions among children of different backgrounds, using heterogeneous groupings. Elliot Aronson's jigsaw classroom, developed around Austin TX in the early 1970s to facilitate racial integration, is notable in this regard. A jigsaw lesson is one in which students, as members of base groups assigned a task with multiple discrete components, each contribute distinct elements to the larger task, fitting the pieces together like a puzzle. Prior to convening the base groups, students master their separate tasks individually or in distinct sets of work groups. Jigsaw lessons thus emphasize interdependence among students.[3]

While factors like race and gender may be considered in designing a cooperative lesson in philosophy, for the purposes of this article, I will mainly consider groupings only according to student abilities. I will not specify measures of these abilities, leaving them to the instructor.

There are three types of groupings of students by ability: heterogeneous, homogeneous, and random. Heterogeneous groupings are those in which each group has students with a range of abilities. Most obviously, one can group academically strong, middle, and weak students together. Less obviously, one can group students according to their ability to negotiate social interactions, distributing the most gregarious and the most reticent students among the different groups. Note that to form heterogeneous groups, the instructor must know the abilities of her/his students in advance. The same holds for forming homogeneous groups. As I will discuss, this is one factor which leads me frequently to dispense with such prescribed groupings.

Here is a simple method for assigning heterogeneous groups. First, determine the desired size of each group. Optimal group sizes depend on the lesson, but are best kept small, since the number of interpersonal interactions in a group grows quickly with the size of the group. The number of groups will be approximately the number of students in the class divided by the desired number of students in each group. (The quotient will often not come out evenly, and the instructor must decide between an extra, small group and a small number of slightly larger groups.) Let's consider an example in which there are eight groups, of three students each, in a class of twenty-four students. To form heterogeneous groups, assign each of the eight strongest students (however one defines 'strongest') to a different group and each of the eight weakest students to a different group. Lastly, distribute the remaining eight, medium-ability students to different groups. Each group of three will then have a strong, medium, and weak student. These group assignments are best done before class, so that when class begins, the group assignments are easily announced.

Heterogeneous groupings facilitate productivity throughout the class. Consider a cooperative-learning exercise which involves group discussions of students' pre-theoretic views on the moral permissibility of torture. Such a lesson would require extensive and delicate interpersonal interactions. It would be helpful to have a student in each group who could start discussions and elicit participation from other students without letting the discussion become too highly charged. In this lesson, grouping students heterogeneously according to their social skills could be useful. In a more technical exercise, e.g. a collaborative analysis of Gettier-style cases, groups will be more productive if each one has at least one student who has a good grasp of the questions going in. In such a lesson, grouping students heterogeneously according to their philosophical abilities may be more advantageous.

Still, heterogeneous groupings suffer from some significant failings. Stronger students can become frustrated with weaker students, and take over the work. When faced with a task to be completed in a short class period, stronger students often become impatient: "Class is almost over. Just let me write it up, OK?" Defenders of heterogeneous groupings point out that stronger students can learn by teaching weaker students. Some teachers even explicitly give some of their stronger students roles like group leader, or student instructor. While stronger students may learn some content by adopting an instructional role, one must sincerely ask whether such students are well-served in this role, or whether such procedures substitute the learning of social skills for course content. Strong students in heterogeneous groups can feel slighted by the undeserved pressure to adopt a pedagogical role, or to carry the brunt of the group's work. Indeed, I had to overcome memories of my own such feelings when I started using cooperative learning in my classes.

Conversely, weaker students, when grouped with stronger students, often become frustrated with themselves and embarrassed, and, feeling that they have little to contribute, turn off. In classes with heterogeneous groups, I always find some students feeling excluded from their groups, having little to contribute. In a lecture class, weaker students can quietly pay attention and take notes. In a cooperative-learning exercise, their weaknesses are made evident, to themselves and their peers.[4]

Homogeneous groupings, in which students in each group are evenly matched, avoid some of the difficulties of heterogeneous groupings. To assign homogeneous groups of size n, assign the n strongest students to one group, the next n strongest students to another group, and so on until all of the students are grouped. Again, one may organize according to students' strengths in various areas including philosophical ability and social adeptness. Again, group assignments are best determined beforehand, and merely announced in class.

Assigning groups of homogeneous ability avoids the problem of having some students naturally take over a group, excluding their weaker peers. Groups of stronger students are usually highly productive. Students in weaker groups find that they can not rely on someone else to do the work and are forced to get involved. The instructor, moving around the classroom, can easily work with many of her/his weakest students at the same time.

Assigning groups of homogeneous inter-personal skills can also protect the quieter student, who is not stampeded by more extroverted partners. I have encouraged boisterous groups to work elsewhere, in a nearby lounge or empty room, in order to reduce the noise that inevitably arises in the classroom during a cooperative-learning exercise. Indeed, I usually extend the offer to work elsewhere to all groups, and find that some students feel empowered by the freedom to work independently for a while.

While homogeneous groupings avoid some of the problems of heterogeneous groupings, they lead to other difficulties. I have found it tempting, especially in classes with homogeneous groups, to admire the work done by stronger students. Such groups can serve as models to the other students. For some students, watching the stronger students engaging in serious debate, or working out a difficult problem, can be as exciting as it is for the instructor. But, I have found that many students are intimidated by their stronger peers, and dismiss their work as irrelevant to their own.

Weaker groups have difficulty getting started on an assigned task. Low-achieving groups tend to drop out of the lesson. The instructor who wants to circulate and work a bit with all the students finds her/himself spending most of the class time with the weaker groups, ignoring the stronger students who could benefit from further challenges. Also, if the tasks for each group are roughly the same level of difficulty, the stronger groups will complete their work quickly, while the weaker groups struggle, and lag. Often, a weaker group will abandon all hope of completing an assigned task, and attempt to use the time for remedial instruction.

Furthermore, as I mentioned, for the instructor to establish either homogeneous or heterogeneous groups, s/he must have a pretty good sense of the students in advance. If s/he wants to conduct a cooperative-learning exercise early in the term, or in a large class, the instructor generally lacks relevant information about the students. Such a deficit is especially problematic if one wants to establish long-term groups, for projects like setting up a course wiki or blog. If groups are assigned early in the term, they are likely to need re-evaluation along the way, disrupting the often-delicate social arrangements that have already been established among the students.

Both heterogeneous and homogeneous groupings have advantages. They allow the instructor to manipulate groups to ensure a desired distribution, which can be useful, depending on the goals of the instructor and the specific assignment.[5] In contrast, random groupings, which are often the default option for instructors who do not take the time to prepare groupings of either other sort, seem to have little besides spontaneity in their favor. If students are grouped randomly, some groups will be more homogeneous and some will be more homogeneous. The instructor of a cooperative lesson with random groupings is faced with a confusing assortment of group types and must manage the problems of both kinds of groups, without being able to ensure the benefits of either.

Still, both heterogeneous and homogeneous groupings have significant disadvantages, too. Furthermore, there are factors favoring the method in which random groupings can be assigned that often outweigh concerns about the mixture of group types. I discuss these factors in the next section.

Assignment methods

Call a method for partitioning a class top-down when an instructor assigns students to groups. Call a method for partitioning bottom-up when students choose their own groups. Lastly, call a method independent when groups are chosen neither by students nor an instructor. Independent methods naturally lead to random groupings, though random groupings can be made by top-down, bottom-up, or independent methods.[6] In contrast, despite the differences between heterogeneous and homogeneous groupings, both are normally the result of top-down assignments.

There are many independent methods of grouping. The easiest one involves counting-off by the number of students in the class divided by the number in each resulting group. For example, if we wanted three-membered groups in a class of twenty-four students, we would count-off by eights. After counting-off, all the ones form a group, all the twos, all the threes, etc. Classes of twenty-five and twenty-six students can also count by eights; one or two groups will have four members.[7]

It is important to note that counting-off is sharply different from having students form groups according to whom is sitting nearby (e.g. “Just turn your desk to the person next to you”) unless the seating assignments themselves are independent. Students often sit with their friends, so forming groups with the students closest to them is likely to lead to a bottom-up grouping.

For a cooperative lesson using the *Objections and Replies* to Descartes’s *Meditations*, I use the following independent method of assigning students to groups. I have printed and laminated pictures of Descartes and his objectors. I cut some of these pictures into three pieces, and two of them into four pieces (for classes with a number of students not divisible by three). In class, I mix the pieces into a basket which I have the students pass around, each drawing a piece, while I introduce the lesson. Groups are formed by the students finding the puzzle pieces that fit with their own. Students find their groups quickly, and they tend to enjoy the short exercise. I use a similar technique for a logic lesson, using pictures of logicians.

In the remainder of this section, I argue that independent methods of assigning groups, especially for single-lesson cooperative-learning exercises, are often preferable to both top-down and bottom-up methods. In many cases, the advantages of independent methods outweigh the advantages of either homogenous or heterogeneous groupings, given the problems of top-down methods that I will discuss. My evidence is mainly anecdotal, from twenty years of using cooperative learning in high schools, colleges, and universities.

While the instructor in any cooperative-learning lesson is focused on organizing the class and on the content of the lesson, students are also anxious about interpersonal social issues. To many students, especially at the start of a cooperative lesson, the composition of their group seems as important as the work they will be doing. Cooperative lessons can bring out problems of social hierarchies and cliques, as students are required to interact with each other. As a teacher, I tend to be nearly oblivious to all but the most obvious manifestations of the social strata in my classroom, to who is the star athlete, the cool musician, the geek. I often have little idea who will be eager or reluctant to work with whom. In contrast, I have found that students are adept at signaling their differences to each other. Even in classes in which students do not know each other at the beginning of the term, their clothes, the ways they talk in class and ask questions, and where and how they sit reveal a lot to one other.

When students are grouped by top-down methods, they often spend a bit of distracted time wondering about the reasons for their placements: “Why am I in this group, rather than that one, with these people rather than those?" Students placed in weak homogeneous groups tend to realize this demoralizing fact quickly. I have seen strong students in heterogeneous groups express disappointment when groups are revealed, anticipating a waste of their time. When I have used top-down methods, I have had students ask to change groups; now I dissuade such requests in advance: "These will be your groups; please do not ask for a change."

Some students interpret top-down group assignments, especially if they are not grouped with friends, as punishments or manipulations. These problems are not ameliorated by top-down random groupings, since students still wonder why they are being grouped the way that they are. These distractions lead to an unproductive beginning to any group activity. Further, and practically, both heterogeneous and homogeneous groupings require advance planning; classes with any level of absenteeism require rapid, last-minute adjustments.

Problems of social hierarchies and cliques are even more prominent in bottom-up group assignments. If left to choose their own groups, students will almost invariably choose to work with their closest friends in the class. Some students have lots of friends in the class, others have few friends. The instructor is likely to have to facilitate some groupings: "Jon, why don't you work with Alysha and Noah? And, we'll put Laura with Daniel and Pedro." Placing students in groups is often awkward and embarrassing, especially for the lonely students who have to be added to groups which have quickly formed. For long-term group projects, especially ones for which the students must meet outside of the classroom when the instructor is not available to facilitate interactions, bottom-up methods may be useful. For in-class exercises, especially ones in which time is short, I find that bottom-up methods are difficult to manage smoothly.[8]

Independent methods for random groupings avoid having students wondering about manipulation from the instructor: every student starts the task on equal ground. Independent group assignments made transparently in class presume and display no preference among students, and can minimize the harmful effects of students' concerns about how groups are formed. I have never had students ask to change their groups when they are assigned independently, though I usually only use independent methods for single-class exercises.

Working with the random groups that tend to result from independent methods of assigning groups requires a bit more agility on the part of the instructor than working with uniformly homogeneous or heterogeneous groups. Classes run using cooperative-learning techniques tend to be active and noisy, and it is impossible for the instructor to follow all of the interactions. When using a cooperative lesson, I pick particular traits on which to focus, when walking around the classroom. If each group has at least one academically strong student, I focus on facilitating social interactions, looking for students who are not facing their other group members, or who are reading instead of interacting with their groups. I ask the more reticent students to take social roles within the group, like note-taker. In contrast, if each group has at least one socially adept student, I pay closer attention to the content of the groups' interactions. When I know that there will be homogeneously strong and weak groups, I prepare assignments with different levels of difficulty, or I prepare extra questions for the strong groups. With random groups, the instructor must be prepared to perform any and all of those tasks, and to glide smoothly among them.

In any cooperative-learning lesson, some students will work well with others, and some students will not. Some groups will be more productive than others. For long-term group assignments, exercises which span several classes or the whole term, the instructor might be well-served to assign groups, or even to let students choose their own, as the advantages of the instructor's reflective intervention or the students' desires to work with their friends might outweigh the disadvantages I have mentioned.

Conclusion

My claims in this paper, especially those in defense of independent methods of assigning students to groups, draw in large part on my experiences with a variety of cooperative-learning exercises. Most research on cooperative learning is done in elementary and secondary classrooms. I have found that the results from younger classes typically transfer quite naturally to undergraduate university classrooms. But more research, especially in philosophy classes, and on interactions between social dynamics and group composition, would be welcome.

Works Cited

Aronson, Elliot, Nancy Blaney, Cookie Stephin, Jev Sikes, and Matthew Snapp. 1978. *The Jigsaw Classroom*. Beverly Hills: Sage Publications.

Aronson, Elliot. 2000-20009. *Jigsaw Classroom* website. http://www.jigsaw.org/

Johnson, David W. and Roger T. Johnson. 1985. "The Internal Dynamics of Cooperative Learning Groups." In Slavin et al. 1985.

Johnson, David W., Roger T. Johnson and Karl A. Smith. 1998. *Active Learning: Cooperation in the College Classroom*. Edina: Interaction Book Company.

Kagan, Spencer. 1992. *Cooperative Learning*. San Juan Capistrano, CA: Kagan Cooperative Learning.

Marcus, Russell. 2010. "A Cooperative-Learning Lesson Using the *Objections and Replies.*" Elsewhere in this newsletter.

Marcus, Russell. In preparation. "A Jigsaw Lesson for First-Order Logic Translations Using Identity."

Marcus, Russell. 1998. "Cooperative Learning on the First Day of Class." *The APA Newsletter on Teaching Philosophy* 97.2: 135-7.

Melser, Nancy Armstrong. 1999. "Gifted students and cooperative learning: A study of grouping strategies." *Roeper Review* 21.4: 315.

Schullery, Nancy M. and Stephen E. Schullery. 2006. "Are Heterogeneous or Homogeneous Groups More Beneficial to Students?" *Journal of Management Education* 30.4: 542-556.

Slavin, Robert. 1995. *Cooperative Learning: Theory and Practice*, 2nd edition. Boston: Allyn & Bacon.

Slavin, Robert, Shlomo Sharan, Spencer Kagan, Rachel Hertz Lazarowitz, Clark Webb, and Richard Schmuck, eds. 1985. *Learning to Cooperate, Cooperating to Learn*. New York: Plenum Press.

Endnotes

1. 'Cooperative learning' has become a term of art in education, referring to exercises in which a class is partitioned into groups. By using the term, I do not mean to imply that traditional lecture-and-discussion classrooms are uncooperative, nor that students engaged in cooperative-learning exercises will always be cooperative!

2. See Marcus 1998, 2010, and In Preparation.

3. See Aronson et al. 1978 and Aronson 2000-9. For a jigsaw lesson in symbolic logic, see Marcus, In preparation.

4. In contrast to some of my experience, Johnson and Johnson 1985's broad survey of cooperative, competitive, and individual styles concludes, "There can be little doubt that the low- and medium-ability students, especially, benefit from working collaboratively with peers from the full range of ability differences. There is also evidence that the high-ability students are better off academically when they collaborate with medium- and low-ability peers than when they work alone; at the worst, it may be argued that high-ability students are not hurt by interacting collaboratively with their medium- and low-ability classmates" (p 118). Some of the evidence in their survey came from college classes.

5. Schullery and Schullery 2006 found positive outcomes associated with both group heterogeneity and group homogeneity, but they were controlling for personality types, using analogs of Meyer-Briggs Type Indicators, rather than for academic ability levels. Heterogeneous groups correlated with student reports of improvements in speaking up, arguing a point, and organizing and presenting thoughts (for females), among others. Homogeneous groups were correlated with student reports of improvements in shyness among males, and with higher grades.

6. Actually, bottom-up methods for assigning groups may undermine an intended random grouping, since students tend to gravitate toward peers of similar ability levels..

7. See Johnson, Johnson and Smith 1998: 2-8 for other interesting independent methods of assigning members to groups.

8. Johnson, Johnson, and Smith 1998 also discourage bottom-up assignments. “The least recommended procedure is to have students select their own groups” (p 2:10). As justifications, they cite a tendency toward homogeneity (in ability and race), as well as an increase in off-task behavior.