



1997

Minding the Mental: Intentionality, Consciousness, and Daniel Dennett in Contemporary Philosophy of Mind

Matthew T. Dusek '97

Illinois Wesleyan University

Recommended Citation

Dusek '97, Matthew T., "Minding the Mental: Intentionality, Consciousness, and Daniel Dennett in Contemporary Philosophy of Mind" (1997). *Honors Projects*. Paper 8.
http://digitalcommons.iwu.edu/phil_honproj/8

This Article is brought to you for free and open access by The Ames Library, the Andrew W. Mellon Center for Curricular and Faculty Development, the Office of the Provost and the Office of the President. It has been accepted for inclusion in Digital Commons @ IWU by the faculty at Illinois Wesleyan University. For more information, please contact digitalcommons@iwu.edu.

©Copyright is owned by the author of this document.

Minding the Mental

*Intentionality, Consciousness, and Daniel Dennett in
Contemporary Philosophy of Mind*

**Illinois Wesleyan University
Department of Philosophy
Research Honors Thesis
Spring, 1997**

MATTHEW T. DUSEK

Minding the Mental

*Intentionality, Consciousness, and Daniel Dennett in
Contemporary Philosophy of Mind*

Matthew T. Dusek

“In coming to understand anything we are rejecting the facts as they are for us in favour of the facts as they are.”

“We should never ask of anything ‘Is it real?,’ for everything is real. The proper question is ‘A real *what?*,’ e.g., a real snake or real *delirium tremens?*”

– C.S. Lewis

AN ILLINOIS WESLEYAN BOOK
Published by The Philosophical Underground

to

Kevin Lannon

philosopher incognito

CONTENTS

<i>Preface</i>	8
<i>Introduction</i>	10
PART ONE—Intentionality	
<i>A Critique of the Intentional Regress Objection</i>	17
(i.) Intentionality and Two Problems of Intentionality	17
(ii.) The Intentional Stance	21
(iii.) The Intentional Regress Objection	26
(iv.) The Regress Objection Rebuffed	28
(v.) Concluding Remarks	31
PART TWO—Consciousness	
<i>Fencing with Jackson and Dennett (What Mary Could Not Know)</i>	34
(i.) Jackson’s Lunge—“The Knowledge Argument for Qualia”	34
(ii.) Dennett’s Parry—“The Illusion of Imagination”	37
(iii.) Dusek’s Parry—“A Myth of Physical Information”	40
(iv.) Strategic Speculation	42
(v.) Dusek’s Riposte—“A Blunt Point”	45
APPENDIX A	
<i>Power, Benevolence, and Suffering: The Moral Obligation Not to Procreate</i>	48
(i.) The Suffering Argument for Atheism	49
(ii.) The Principle of Benevolent Action Abstracted	50
(iii.) The Benevolence Argument Against Procreation	54
(iv.) Criticisms and Their Responses	56
APPENDIX B	
<i>Logic of “The Benevolence Argument Against Procreation”—Symbolic Notation</i>	62
<i>Bibliography</i>	63

PREFACE

This thin volume, which I have presumed to title *Minding the Mental*, represents the culmination of my research honors project in philosophy at Illinois Wesleyan University. The project began in the Spring semester of 1996 as a continuation of my interests in the philosophy of language, which had bloomed under the tutelage of (the very Fregean) Professor L.W. Colter. Professor Leonard Clapp graciously agreed to act as my thesis advisor, and in addition to the papers which we read together I attended a significant portion of his enjoyable course on the philosophy of language.

Approximately half-way through that semester, the focus of the project underwent a shift, largely as a result of reading and discussing some influential papers by Hilary Putnam. In response to certain problems of intentionality which arose, Professor Clapp and I turned to Daniel C. Dennett's *The Intentional Stance*. Through mutual interest and consent, we opted to make the research honors project a project in the philosophy of mind, with emphasis on the writings of Daniel Dennett.

Following the work on intentionality, during the May short-term of 1996, we looked into problems of subjectivity and objectivity and their treatment by the philosopher Thomas Nagel. Those readings would prove valuable in understanding some of Dennett's theoretical and methodological commitments in both his theory of intentionality and his theory of consciousness.

We read about Dennett's theory of consciousness, and the difficulties presented it by Frank Jackson's famous "Knowledge Argument for Qualia," in the early portion of the Fall semester. This work was temporarily interrupted and delayed by the development of a paper, "Power, Benevolence, and Suffering: The Moral Obligation Not to Procreate" (included as Appendix A of this text), and returned to in the latter portion of the semester.

The task of compiling the bibliography and preparing the thesis manuscript, as well as further preparatory reading, has been completed for formal defense during the Spring semester of 1997.

Thanks where thanks are due.

My research advisor and sometimes mentor Lenny Clapp, more than any other, has been a wonderfully instructive guide and guardian along the way. I know that if I have succeeded to some degree in adopting his *mantra* of clarity, logical rigor, and the tireless interrogation of meaning, I am a better thinker as a result. However, I have benefited greatly from several other individuals as well, both in and out of the classroom; I would be remiss not to acknowledge that debt.

Preface

Thanks go to Charlotte Brown, Larry Colter, Pat Francken, Jeff Fry, Carl Gillett, Ted McNair, and Deb Waldman. I regard them as teachers and friends, all. Although not a part of my formal philosophical training, Professor William White was a critically valuable influence in the decisions which led to it. And, as ever, I am most grateful to Rachel Gulling and Kevin Lannon, who serve as a constant source of criticism, inspiration, and friendship. Finally, to all who have read drafts of earlier versions of this project or discussed with me the issues around which it centers, I am grateful. The final product has been shaped, in ways subtle or obvious, by each piece of input.

A special thanks to the members of my research honors review committee, for reading yet another verbose text *à la Matt* and kindly agreeing to collectively drill me on the positions endorsed: Leonard Clapp, convenor (philosophy); Charlotte Brown, department chair (philosophy); Carl Gillett (philosophy); and Robert Bray (English).

It is my sincere hope that all involved will find it worth the effort.

MATTHEW T. DUSEK
Illinois Wesleyan University
Department of Philosophy
April 1997

INTRODUCTION

“What is the answer? . . . In that case, what is the question?”

—Gertrude Stein

The mind. *Sanctum sanctorum* of subjectivity. Soundstage of the mental. Consciousness’ cockpit. Romping-grounds of the intentional. A great deal, it would seem, rides on the notion of *mind*. It’s not just that naughty children never do, or that people when irritated often claim to have half-a-one. Though perhaps telling in other ways, it isn’t so important that while we all think we lose ours from time to time, we rarely—if ever—doubt that we had one to begin with. Solipsists are perfectly willing to doubt that anyone else actually possesses one,¹ but no one suspects that everyone *but* herself has one. The reason seems clear and distinct as Cartesian water; it just makes sense, which is another thing that minds are rumored to be responsible for. One could go on and on—the idioms seem endless—but nevermind.

Why *does* the idea of the mind seem so important? Though it would be a shock, and a major setback to contemporary understanding of anatomy and physiology, one could at least conceivably deal with the startling discovery that the biological apparatus responsible for human cognition is not the brain, but in fact the vertebrate endo-skeleton. Yet, in contrast, it just seems absurd to suggest that we are mistaken in identifying the mind as that which is constituted of and/or responsible for various mental events. “Dammit,” one might say, “what I mean by ‘mind’ just is that which is constituted of and/or responsible for various mental events, whatever it turns out to be.”

Perhaps, then, our minds are important to us because we all have them and we know it. I personally don’t know first-hand whether or not I have a brain—I’ve never seen it—and for that matter I’ll only be taking it on the word of others if I believe that that gooey bit of gray matter upstairs is the thing that ultimately lets me get bothered over skeptical doubts and bewitched by natural language and tangled up in all the other sorts of mischief my thoughts get me into. But our

¹Solipsists have likely spent much time at the dreaded four-way intersection.

Introduction

minds are different. I don't need to open up my skull for confirmation. I know I have a mind, because I have the appropriate thoughts, beliefs, and sensations. As L.W. Colter has observed, "[t]o all appearances, we enjoy (or, in some cases, suffer from) very rich 'inner' lives."² And that's all I need in order to know I have a mind, it seems, because a mind just is the vehicle of those inner lives, isn't it?

But there is also a different sort of thing going on. For, whereas the word 'brain' designates an organ, and that organ is responsible for whatever it is responsible for (things we discover after having fixed the referent of 'brain') independent of the way I use my words, minds aren't like that. For one thing, we have no direct acquaintance with the object which is the nominatum of 'mind'. We only have such acquaintance with some mental phenomena. Not only have I never seen my mind, I don't even know what sort of thing I should be looking for—or where I should be looking—to find it. Is it in my head with my brain? Is it my brain? Is it somewhere else in my body, or maybe distributed throughout my body? Is it somewhere else entirely, in a realm totally unlike the physical realm of extended objects? And what sort of a place would that be? Although most of us don't know by experience that we have brains, brains themselves are like ships and shoes and cabbages and kings. They're concrete, medium-sized objects in the observable world, definite, and when we refer to them, we know what it is to which we're referring. But minds, at best, are on a similar footing to subatomic particles—whose existence cannot be observed but must be inferred—or, at worst, are on similar footing to such *faux amis* as the ether of early physics or the humors of medieval physiology. To be sure, the way we think and talk about the mind enjoys a certain flexibility not shared by the way we think and talk about the brain. In fact, far from being similar to the way we use 'brain', our usage of the term 'mind' appears almost *bass ackwards*. We start with phenomena and posit a special sort of organ to which these phenomena must owe their very existence, or at least their status.

Why this debt, and why the inclination to posit such a mysterious, bizarre, utterly unseen "mechanism"? Well, because all these phenomena share a tauntingly similar characteristic—call it,

²Unpublished manuscript.

the mark of the mental. It is taunting, for what better to explain the mental but minds? And isn't it appropriate that our minds are not observable to others like our brains (potentially) are? My thoughts certainly are not accessible to observers like my behavior is; they're hooked up intimately with *me*. And so must be my mind. And thus, perhaps the idea of the mind seems important because it offers some hope of accounting, in one neat package, for the mystifying phenomena which are connected so personally to my awareness of myself—my thoughts, beliefs, hopes, fears, desires; my sensations; my memories of all or any of these. These things which lie at the “center,” truths too important to be undone by a weather report, may contribute to what is sometimes an almost dogmatic death-grip on our idea of mind. Indeed, then, a great deal does ride on this notion.

But ought it to? Or need it, at any rate?

With the sweeping progress of modern science and the conceptual revolutions which it sometimes engenders, it is becoming less clear that so much actually does. The seed which once belonged almost exclusively to the armchair ruminations of the philosophers has been swept up into scientific pursuits and has flowered tremendously. Now “the mind” is of intense interest not only to philosophers but also to psychologists, computer scientists, and even physicists. There has been extensive research in neuroanatomy, neurophysiology, cognitive psychology, artificial intelligence, and on and on. We are in the midst of what has been called a cognitive revolution.

In the wake of science's increasing ability to provide answers where before there was only mystery, there is fresh incentive to question older modes of thinking about the ways in which we think. In some cases, this leads to new criticisms and challenges, but in others it leads us *back* to those which are perennial. It is here, especially, where the philosopher is likely to make valuable contributions to cognitive science and the project of coming to understand the mind.

For example, philosophers are often concerned with matters such as explanatory efficacy. If we only posit the existence of minds by inference from various “mental” phenomena, by what count is the notion of mind moving us in some way beyond our understanding, or lack thereof, of those phenomena? Particularly now, when research leads to more satisfactory understandings of

Introduction

many of these phenomena, what motivation remains for clinging to our old folk-psychological notions? This question, and others similar to it, is frequently posed by a school of theorists called eliminativists.

However, given the length of time during which philosophers have had opportunity to debate these things, a mature dialectic has developed, one which exposes certain gaps that mitigate the impact of eliminativist arguments. Two areas in particular—rather large gaps, each—have proven stubbornly resilient against the sort of physicalistic analysis common to the scientific enterprise, and these areas constitute the bulk of the work that is done in contemporary philosophy of mind. They are intentionality and consciousness. And thanks in large part to them and the difficulties they present, the mind as an ontologically robust entity rather than merely as a cognitive psychologist’s “level of abstraction” is in some circles a heartily supported thing. With respect to the mind, the members of these circles are familiarly known as realists.

And so there is controversy.

The big issue is: How do our folk psychological notions—intuitions that we really do believe things and that we experience the world in some sort of *phenomenal* way—square with the advances of science and its dominant paradigm *physicalism*? Are they compatible? It often appears they are drawn into diametric opposition. If they are incompatible, which are we to abandon? Folk psychology has succeeded in getting us through the day relatively intact for ages; folk-psychological talk allows us a great range of interpretive and predictive power. Are we to trade this away in favor of a more accurate, scientific model? For, surely, we cannot give *it* up. Science stands as the single-greatest hope for a unified body of human knowledge, possibly the only hope. Or is there perhaps some fundamental undecidability, a necessary schism, which prevents a satisfactory resolution of this problem? Certainly many have thought that there is something peculiar about the self-referential task of thinking about *thinking*. (Though others might see it nevertheless as a very important task for any epistemically responsible community of thinkers.)

The issue is big even for those—especially for those—already committed to a particular camp. For physicalist philosophers of mind, it is decidedly important to affect some sort of

reconciliation—or to firmly establish the unimportance or impossibility of one—between the phenomena of consciousness and intentionality, and the physicalist’s ontology. It is also important to elaborate some sort of coherent theory of each. For reasons which will become apparent further in this thesis, the importance of this task does little to alleviate its complexity. The field is largely divided. And although some have a theory of one or the other, few have developed a theory of both.

Daniel C. Dennett is among those few, and his work has been heavily discussed in the contemporary literature on the subject. Although his naturalistic ambitions are fairly traditional now, when he entered that literature in 1969 with the publication of *Content and Consciousness*, “its resolute naturalism and earnest concern with what science could tell us about the mind struck [him] as quite pioneering—or quite eccentric, depending on [his] mood” (ix). Since then, he has published another seven books, several of which extensively expound upon, or bolster the theoretical underpinnings of, the basic positions already endorsed in his early work. One is a volume on evolutionary theory and morality written in support of his views on intentionality. One is an investigation of the problem of free will, which threatens his materialism. Another is a work on types of minds written in response to the implications of his views on mind. The works principally treated in this thesis are his *The Intentional Stance*—a more mature exposition of his theory of intentionality—and *Consciousness Explained*, in which Dennett argues against the “Cartesian theater” model of consciousness.³

Rather than treating each of these directly—a substantial undertaking—I have opted to forego anything approaching a comprehensive evaluation of his theories, and instead have selected a more humble aim. Criticisms abound in the intellectual world, and in philosophy especially, and Dennett is the academic equivalent of a walking bullseye. In Part I of this thesis, I select an interesting criticism of Dennett’s *intentional stance* (a criticism which I formerly propounded), examine the dialectic involved, and issue a rejoinder. In Part II, I engage in a similar project, this time

³For a witty exposition and defense of the “Cartesian theater” model of qualia and consciousness, see Lormand (1994).

Introduction

choosing Frank Jackson's Knowledge Argument for Qualia as the threat to Dennett's theory of consciousness (a threat which, I feel, Dennett himself fails to defuse).

The final goal of this thesis, *Minding the Mental*, then, is—though critical in method—a constructive one. It shall highlight some important, motivating issues in contemporary philosophy of mind; argue against some persuasive mistakes which cloud those issues; and leave Dennett's theories unfettered, to be evaluated on other grounds. Whether, in that separate arena, they can stand against the lions or not is another matter, one which I do not pretend to take up here.

Part One

INTENTIONALITY

“So what kinds of things have content? Well, if we must speak this horribly misleading way, *people* have content, because *people* have beliefs and desires. What are beliefs? Well, this is a silly question—the person asking it is bewitched. What is content? Well, this is even sillier.”

— Leonard Clapp (*The IWINK Proceedings*, 1996)

A CRITIQUE OF THE INTENTIONAL REGRESS OBJECTION

I believe a number of things.¹ I desire more than a few as well. But does it follow from this that I *have* beliefs and desires, and if so do I have them in the same way that I have a car, several pairs of Levi's, and annoying problems with allergies? Or am I perhaps just "bewitched by natural language" if I think so? It is unclear at this point just what to make of questions such as these and even less clear what to make of the baffling issues around which they orbit. In this paper I shall examine two problems which face contemporary speculation about beliefs, desires, and other sorts of intentionality. Subsequently, I shall evaluate Daniel Dennett's account of intentionality as it is presented in his *The Intentional Stance* (1987) and a compelling criticism of that account. I shall then offer a critique of that criticism. Finally, I shall conclude by recommending an open agnosticism with respect to Dennett's theory.

(i.) *Intentionality and Two Problems of Intentionality*

Ultimately, we hope to come to terms with just what intentionality is, including the particulars regarding what makes something an intentional system and what distinguishes it from a system which is not intentional. At the outset, however, we must at least have a working concept with which to identify our subject. By our common usage, what do we mean when we speak of something as being intentional?

Intentional states are, generally speaking, characterizable as representational states, i.e., states which are *about* something or other. For example, when one believes of *The Lord of the Rings* that it was written by J.R.R. Tolkien, one has a particular intentional state (belief) *about* J.R.R. Tolkien's epic trilogy. Along these lines, believing *that*, desiring *that*, hoping *that*, opining *that*, etc. are all instances of intentionality, and being "in the state of" any one of these is being in a particular intentional state. When we talk about something as being intentional, we mean that it

¹Perhaps not so many now as when I began studying philosophy.

A Critique of the Intentional Regress Objection

is capable of representational states such as belief, desire, hope, opinion, and so on. It is the property of *aboutness*, then, which is central to the notion of intentionality.

As should be immediately obvious, intentionality is not something rare and unobtrusive that the philosophers dreamed up and with which the lay-person has little or no acquaintance. On the contrary, it pervades our lives. We make sense of ourselves, others, and at times even the world through intentional characterizations. Moreover, we always choose our actions on the basis of our beliefs, desires, hopes, and opinions. Yet, if it is true that intentionality is such a familiar part of our lives, in what way is it, as I suggested earlier, “baffling”?

The answer has to do, among other things, with explanatory consistency in our cosmology. Dennett, in his Foreword to Ruth Millikan’s *Language, Thought, and Other Biological Categories* (1984) writes,

... since we human beings are a part of nature—supremely complicated but unprivileged portions of the biosphere—philosophical accounts of our minds, our knowledge, our language must in the end be continuous with, and harmonious with, the natural sciences.

L.W. Colter in an unpublished manuscript spells out explicitly what seems to be implicit in Dennett’s requirement:

Our cosmology tells us that everything in the universe is a function of the Big Bang and the universe’s evolutionary history. Furthermore, the atomic theory of matter tells us that everything in the universe is completely made up of the particles explained in the atomic theory. Hence, human beings, continuous with nature as they are, are completely so made up. Now a complete science of the universe will . . . have as a part of it a complete science of human beings. Thus, a complete science of human beings will be a strictly physicalist science.

It is the difficulty of accounting for intentionality under the stipulations of our (physicalist) cosmology which makes the subject so baffling in contemporary philosophy. We may circumvent the difficulty by abandoning some part of the cosmology, such as the atomic theory

A Critique of the Intentional Regress Objection

of matter, or by denying that human beings are continuous with nature, but these moves are purchased at a high price. In fact, were we to do so, it would suggest that we wished to salvage some sort of privileged account of the mind in a largely arbitrary and unwarranted, even intellectually lazy, manner. The legacy of the scientific revolution, at least for the science of intentionality, is not technological advancement but philosophical puzzlement. We must be forthright in admitting this unfortunate state of affairs and only then attempt to move beyond it.

While I have stated that there is a sort of conflict between the very notion of intentionality and the manner in which it must fit into the contemporary physicalist cosmology, I have yet to explain just why this is so. What obstacle prevents us from offering some sort of reasonable hypothesis about the presence of intentional properties in physical systems? What is it about intentionality that causes it to be so difficult to characterize in the physicalist's vocabulary? The answer to this question has to do with the very feature of intentionality which I have suggested is critical to the whole notion, that of *aboutness*, and our inability to account for it in terms of strictly physical relations.

The fundamental problem is roughly this: Our beliefs, desires, suspicions, hopes, and other intentional states are supposed to be about things. The particles described in the atomic theory of matter, of which we (and everything else) are constituted, on the other hand, are not by most accounts about anything. However, if we are to be fully explained in a way consistent with our cosmology, then this aboutness feature of (our) intentional states must be somehow construable in terms of the accepted particles, forces, and relations of the atomic theory. How this is to be done is a much debated question.

This problem is one of accounting for the (let's call it) relational role which intentional states play in a manner compatible with a physicalist cosmology. There is another question, similar to that raised by the "relational" problem, which arises in treating intentional states as natural (and physically describable) phenomena. It is an ontological question which concerns whether there are *in fact* these things beliefs, desires, and so forth which we may locate somehow in the universe, be they particles in the brain or psychological states or whatever else. Taking for

granted that people do, and will continue to, speak of themselves in terms of beliefs and desires, what are we to say about the physical reality which either does or does not correspond to the way that people talk?

These two problems together typically motivate one of three types of responses from philosophers of mind. One response is to question the sufficiency of the current cosmological account. Given the apparent inability of intentionality to be accounted for in the physicalist's ontology, but taking it to be indispensable and a real feature of the universe, one abandons unqualified allegiance to the cosmology. This route, though it circumvents the particular problems which I have described, amounts to a very expensively purchased liberation; it leaves unexplained (and possibly unexplainable) many other questions regarding not only intentionality and the peculiarity of mental phenomena but also the new cosmology which is to replace the former.

Another response is to affirm that we as human beings are continuous with nature and must be describable in a manner consistent with our cosmology but to deny that there really are such things as beliefs and desires. Philosophers who opt to defend this position are known as eliminativists, for they attempt to explain away our folk-psychological intuitions about belief and desire in terms of a more ontologically conservative neurophysiology. They benefit by escaping the difficulties posed by the ontological problem—such as dealing with Putnam's famous twin-earth problem (1973), but must still account for the relational problem of how physical stuff can somehow manage to be about something. Moreover, there remain certain concerns about how our folk-psychology could manage to be so incredibly successful if patently false.

If one finds the first response unpalatable and eliminativism curiously unsatisfying, one will likely find oneself supporting realism. A realist with respect to intentional states will affirm a sub-personal cognitive psychology which preserves our folk-psychological intuitions. We are continuous with nature, the realist-with-respect-to-intentionality will say, and there really are such things as beliefs and desires which are more than convenient or useful fictions. Of course, this salvaging of our folk notions coupled with a commitment to cosmological consistency leaves

the position open to each of the aforementioned problems. And it has been no small matter, attempting to extricate realism from these pitfalls.

All things considered, the range of options seems unsettling at best. It is not just that the appropriate sciences have not *yet* proffered an account of intentionality which confirms the realist's or the eliminativist's picture but are well on their way to doing so; there is reason to at least suspect that they never will be capable of doing so. At this point one might stoically opt to plunge headlong into the task of extricating the theory of choice from the quagmire of the problems above. One might, however, look for other (less orthodox?) alternatives and discover Daniel Dennett and his "intentional stance."

(ii.) The Intentional Stance

Dennett thinks it worthwhile to consider seriously why we might wish to talk about beliefs at all. Not only do I agree with him that it is worth pondering, I think it an excellent way to begin to understand just how his take on a "naturalized" intentionality fits together and to see how it might differ from more traditional approaches. So, why might we wish to talk about beliefs?

The knee-jerk reaction is probably the response that we wish to talk about them because we have them and they are *there* (wherever there is). This response is fairly clearly unsatisfactory, of course, in that it rather ungraciously begs the question against the eliminativists, who feel that they have good reason to doubt that this is the case and good scientific support to back those suspicions. Look anywhere you like inside the head, the eliminativist will tell you, and you're not going to find a belief. You'll find synapses and neurons and so on, but you'll be hard-pressed to locate a representational state anywhere. Yet we certainly do talk about beliefs. What else, then, could suggest itself as a promising answer?

A familiar and plausible reason we like to talk about beliefs and desires and such, Dennett suggests, is that we want to explain and predict human behavior. Positing intentionality is a remarkably successful means of doing this. For example, suppose that I am observed every Saturday afternoon climbing into my car and driving to my Corvette specialist of choice, D & M

A Critique of the Intentional Regress Objection

Corvettes. Furthermore, it is observed that I am prone to spending large portions of my afternoon gazing wistfully at the various well-maintained models in their showroom, lingering longest at the ZR-1. Were one to attempt to explain my behavior, it would be perfectly natural (and completely correct) to speculate that I *desire* to lose myself in the Elysian bliss of Corvette-browsing and *believe* that a trip to Ogden Avenue's slice of automotive heaven will satisfy this desire. Thus, I act on the basis of this belief. Furthermore, one might similarly predict my behavior on future occasions on the basis of these posited intentional states, for instance that next Saturday afternoon I will drive to D & M Corvettes or even that during the summer I will voyage to the Illinois State Fairgrounds for the Bloomington Gold Corvette Show and to the National Corvette Museum in Bowling Green, Kentucky.

The explanatory and predictive power of this strategy is attested to by each of us by our actions many times daily. We constantly and even without conscious consideration successfully predict and describe the behavior of those around us by employing the notion of belief. Nevertheless, a sound account of intentionality must offer more than this. Such a strategy of belief attribution may be successful in the event that there are no such things as beliefs, but it will be at least equally successful if there happen to be such things, as well. So, although this may be the reason that we talk about beliefs, desires, and other intentional states, it cannot purport to tell the whole story. What are we to say, then, when we wish to move beyond useful fictions to a model which reflects a higher degree of competence and understanding? Are there beliefs or aren't there, really? And if there are, how can they manage to be *about* something?

Dennett unfortunately (though perhaps to his credit) does not fit neatly into the categories described above for responding to questions such as these, and he has a rather unique, albeit sometimes frustrating, take on this issue. If one were to inquire of him whether beliefs are real, he would likely reply in all earnestness: yes and no. Of course, an answer such as this requires some explication. It is such an explication which I shall attempt to briefly provide.

Dennett relies heavily in his conception of intentionality on a very strong instrumentalism. In order to make this conception more comprehensible, it will be beneficial to understand

something about Dennett's brand of instrumentalism. As an instrumentalist, Dennett is cautious in talking about the reality of things in the usual unqualified way that people tend to do. Generally, when we wonder whether something is real, we wonder whether that thing is out there in the world in some way or another (even if it is forever beyond the prospects of human observation). And, when we decide that there really is no such thing, we mean more or less that it is not able to be located anywhere in the universe, in any fashion, in practice by human beings or even in theory. But Dennett is uncomfortable with these standard usages.

By Dennett's reckoning, a thing is "real" or really there just in case it makes sense to talk about it as if it is. For example, suppose that we view a computer running John Horton Conway's Game of Life.² For now, suffice it to say that the Game of Life will appear as a two-dimensional graphic presentation on a computer monitor. This presentation occurs on a grid, the intersection points of which are in either an ON or an OFF state which varies according to the states of the intersection points on all sides of it. The ON/OFF states may be marked by differently colored pixels or stronger or weaker light intensities or whatever else the computer is capable of. What emerges is a flowing pattern in a nicely Heraclitic state of flux.

Now suppose that you and I both observe the pattern which is currently developing in the Game of Life. Both of us being mathematically-minded people, we quickly pick up on the resemblance of the pattern to mappings of certain mathematical functions, though—interestingly—not the same ones. I remark, "Oh, it's computing the function x " (let x stand for some unspecified mathematical function). But you retort, "No, it's computing the function z " (let z stand for some unspecified mathematical function different from x). Now, it just so happens that the mapping which the Game of Life presents is in fact the very same which we would behold in a mapping of the functions x and z . We are both, it would seem, correct. Yet while in a sense the Game of Life really is computing these functions, it is nevertheless doing so merely in virtue of the fact that the parameters set by the Game of Life caused it to do so. It may be said properly to be computing these functions only because it is *interpretable* as computing them.

²See Gardner (1970).

The same thing goes for entities, not just for processes or states, Dennett thinks. The center of gravity of an object is a real thing, in a sense, because as a theoretical entity it makes sense to talk as if it is. Such instrumentalism is certainly a sort of realism, but it is a sort that must be carefully distinguished from the garden variety. With this brief understanding of Dennett's instrumentalism in mind, we may turn now to his theory proper.

From the outset, it is important to set a firm theoretical foundation. The perspective from which we must assess beliefs and intentional states (as well as nearly any other phenomena), Dennett thinks, is that of the detached and objective third-person. It is unnecessary and even dangerous to appeal to a subjective first-person perspective, for we are liable to be tempted into supposing that we are examining some ontologically robust thing to which we somehow have access. I have already suggested that such a supposition is premature. (Recall that the eliminativist holds such a view of beliefs under high suspicion.)

From this objective vantage point we may utilize certain intellectual tools in order to predict and explain a vast array of phenomena. These tools are what are known as interpretive strategies or predictive stances, and Dennett thinks that there are three of them of note: the physical stance, the design stance, and the intentional stance. Each of these has its own conditions of effectiveness; each excels in explanations and predictions where the others falter or become hopelessly entrenched in technical details.

An example may serve best to illustrate the application of each strategy. Consider a game of chess in which I compete against a chess-playing computer. Quite clearly, this is a case in which it would be desirable to be able to explain and especially to predict the computer's actions. My victory is likely to hinge upon my being able to do so. How, then, might I go about this? For starters, I might try the physical stance. According to Dennett (1987), if I wanted to predict a system this way I would "determine its physical constitution (perhaps all the way down to the microphysical level) and the physical nature of the impingements upon it, and use . . . the laws of physics to predict the outcome for any input."³ This rather impractical Laplacean strategy

³page 16.

requires a great deal of raw computational power and is somewhat obviously inappropriate in this instance. I might turn next to the design stance. From the design stance I ignore the technical details of an object's physical constitution and "on the assumption that it has a certain design, [predict] that it will behave *as it is designed to behave* under various circumstances."⁴ This stance is ordinarily effective when dealing with computers and machines which perform certain low-level tasks but requires just too much detailed knowledge of the program involved to be efficient in this instance. The computer is "designed" to win the chess game by making strategically sound moves. I need to predict what those moves will be and to do so fairly quickly. I may do this by employing the intentional stance.

Using the intentional stance, I decide to treat the computer as an ideally rational agent. Then, given its place in the world and its purpose (making strategically sound chess moves) I figure out what beliefs it ought to have. These beliefs would encompass how particular pieces may be moved and how such moves may be sequenced best. Then, on the same considerations, I figure out what desires it ought to have (the desire to capture my power-pieces, the desire to prevent me from moving my pawns across the board, and so on). Finally, I predict that the computer will "act to further its goals in light of its beliefs."⁵

It seems reasonable to suppose that Dennett is right that we may so utilize the intentional stance to predict the computer's actions. Dennett, however, thinks not just that we may explain the computer if we treat it *as if* it has beliefs and desires and is ideally rational; he thinks that if we may explain it in such a way—i.e., if it is interpretable from the intentional stance—then in a very real way it *does* have beliefs and desires and is intentional. For Dennett all it is to *be* intentional is to be so interpretable. One may be tempted to object that we know already that computers don't have these intentional states and that positing them for the sake of explanation and treating them as real components of its silicon "brain" are two very distinct things. But they aren't for Dennett. As far as he is concerned, all there really is to being intentional—for us,

⁴page 17.

⁵*ibid.*

computers, or anything else—is to be interpretable from the intentional stance. A belief is real, Dennett would say, but only in the same way in which the center-of-gravity of a physical object is real. If we are tempted to say otherwise about ourselves, it is merely because we are more complex from the viewpoint of another stance and have bought into the errors of a first-person point of view.

Why, then, is Dennett's answer to the question *Are beliefs real?* both yes and no? It is no because beliefs are not these things to be found inside heads or corresponding isomorphically to brain states. If this is what is meant by "real" then the answer is most definitely no. But there is another very important instrumentalistic way, Dennett thinks, in which something may be regarded as real. If something is predictable/explainable/interpretable as having a particular property, then we may consider it to actually have that property. We must simply avoid the mistake of thinking the property to be some kind of concrete thing. In this way the answer is yes. Looked at in this manner, we can see that Dennett is a sort of peacemaker or arbitrator between the realists and the eliminativists.

(iii.) The Intentional Regress Objection⁶

So far I have provided a brief summary of Dennett's project in *The Intentional Stance*. I want to turn now to a criticism of that project.

I have explained already that central to Dennett's picture of the mind and intentionality is the assumption that the objective third-person point of view can capture all which needs to be said regarding beliefs and desires. There is a temptation to posit too much about the nature of intentional states when we move from the third-person objective to the first-person subjective point of view. We tend to overlook, in Dennett's opinion, that all we in fact are doing when we speak confidently of our own mentalistic internal states is applying the intentional stance to ourselves. We are applying, it would seem, a predictive strategy to ourselves much as if we

⁶See Kim (1996) for a proponent of this objection.

interpreted our behavior from the design stance or the physical stance. Allowing that all of this is true, will it be enough?

Initially, at any rate, it seems plausible; Dennett is even able to save the realist leanings of our folk-psychological notions in spite of the threat to them that is posed by the relevant sciences. In many ways, Dennett is rather clearly correct about the intentional stance and the ways in which we apply and utilize it. For one thing, it definitely seems that we *do* apply and utilize it (to some degree). We need but recall the example of the chess-playing computer or that of my regular weekend *rendez-vous* with D & M Corvettes to reassure us of this much. Yet one might have doubts that the intentional stance, as Dennett conceives it, can prove ultimately tenable. There is a compelling line of reasoning which concludes that it will not. If it is correct, Dennett's conclusion that intentionality *just is* interpretability under the intentional stance faces a potentially devastating difficulty, one which emerges from his commitment to the objectivity of a third-person point of view. The failure of the intentional stance under this commitment is that it seems to depend in the last analysis upon a regress, in this case a regress of intentional dependence.

The line of reasoning which leads to such a conclusion might go approximately like this:

You may only rightly be said to have beliefs, desires, etc. (intentionality) if I, or someone else, interprets you from the intentional stance. But we mustn't overlook that this act of interpretation is *itself* an intentionally saturated act. I must myself believe things, opine things, suspect things, and so forth in order to interpret you as doing the same. I believe that you have acted in a particular way or are going to act in some way, desire to explain or predict that action, suspect that by positing beliefs and desires in you and treating you as a rational agent I may do so, etc.. Yet, by the stipulations just set forth, I may be rightly said to be intentional in these ways only if I am further interpreted by another in the relevant respect. And that person, in order to interpret me, must be herself intentional; this requires pushing it back a step further. Each case of intentionality seems to depend on the intentionality of a case before it, *ad infinitum*.

If correct, this leaves the intentional stance as Dennett has laid it out in a bit of a conundrum. The series must either recede in an infinite and ungrounded regress, circle back upon itself, or terminate in an independently intentional observer/interpreter—or so it would seem.⁷ Only the last will preserve intentionality (the other two options are incapable of beginning the series), but it will do so at the price of forsaking Dennett's position. Dennett, remember, thinks that all intentionality *is* interpretability under the intentional stance; independent intentionality is clearly contrary to this. It might be thought that we may circumvent this difficulty by maintaining that *we* interpret *ourselves* as intentional, but this will not do either. For that interpretation depends upon our intentionality which depends upon our prior interpretation which itself depends upon our prior intentionality. And, in fact, we are arguing in a circle. The point of this criticism is that, granting Dennett his picture, the chain of interpretation-dependent intentionality must have begun somewhere, yet to maintain that it begins somewhere is to maintain that it is interpretation-*independent* after all.

This objection to the intentional stance, if correct, is quite devastating to the project as a whole; if correct, it successfully undermines the very foundation upon which Dennett's account is built. Nevertheless, an objection stated is quite distinct from a successful objection. I submit that it is far from clear that the intentional stance has been dealt a mortal blow by the intentional regress objection. As I shall proceed to elaborate, the objection takes for granted the insignificance of what turns out to be a rather critical distinction.

(iv.) The Regress Objection Rebuffed

Stuart Watt recently wrote, in a paper entitled "Naive Psychology and the Inverted Turing Test,"⁸

Another possible objection to the inverted Turing test is that it is implicitly recursive; it defines systems indirectly in terms of the ability to recognise their

⁷For a lucid discussion of methods of dealing with putative regresses, see BonJour (1985), where he argues in favor of a modified Coherentist solution to the epistemological regress problem.

⁸PSYCOLOQUY, 12 May 1996

A Critique of the Intentional Regress Objection

own behaviour. This is true; the test is indeed recursive, but it is not an infinite regress kind of recursion, but a transactional, or temporal regress, kind of recursion, going back through social and biological evolution rather than through logical forms. Because the participants are playing transactional games, guessing at one another's mental states, there is an inevitable mutual recursion between their actions. The point to note is that exactly the same is happening both in the standard Turing test and in normal human interaction in the real world.

It is not particularly relevant to the topic at hand what is the background to this statement, nor whether it is correct in its application to the topic of inverted Turing tests. It is suggestive, however, of the possibility that the intentional regress objection presumes too much.

First of all, not all regresses are vicious. It is conceivable that a given chain (of causes, of reasons, etc.) may circle back upon itself or recede indefinitely without becoming threatening, or vicious, to itself. The question, then, arises as to whether the regress suggested might be of an innocuous and non-vicious variety. The answer to that question will depend upon our sensitivity to the aforementioned critical distinction. By the first reading it will turn out that the regress is indeed self-threatening. By the second reading, however, it appears that Dennett is on somewhat safer ground.

For the objection to succeed, any instance of intentionality must rely on some prior instance of intentionality, as intentionality depends upon interpretation and interpretation is itself an intentionally-saturated act. Thus if, as the objection states, you may be considered to be intentional only if I *interpret* you as being intentional, the regress successfully begins and thereby threatens the success of the project. Yet, there is a slip, albeit ever so slight, which purchases this objection the force which it contains.

According to Dennett, for something to be intentional, it must be *interpretable* from the intentional stance. But note that this is different from what the objection requires. For the objection to succeed, that thing must be *interpreted* from the intentional stance. Yet, of course, interpretability and being the subject of interpretation are distinct things, though of course they are related. In the former case (interpretability), it merely needs to be possible for one to be

interpreted. In the latter case, it is not just that it is possible but that it is actually the case that one *is* interpreted. This may seem like splitting hairs, but the value in half a hair becomes readily apparent when we consider the difficulty of trying to frame the intentional regress objection in such a way that it captures not just those interpreted from the intentional stance but those interpretable from it as well.

Consider the story for your intentionality, now. In order for you to be considered intentional, you need to be *interpretable* from the intentional stance. You, in point of fact, are interpretable from the intentional stance, as evidenced by the fact that I interpret you on some unspecified occasion from that stance and successfully explain and predict your behavior. It need go no further; the regress stops here.

A determined objector might opt to push her case. Granted that we must be careful to confine ourselves to interpretability under the intentional stance, she could object, there still seems to be something peculiar going on. For observe, you are intentional just in case you are interpretable from the intentional stance. Purportedly, the story ends here. Yet, just how is it that we are to know that something (you, in this case) is interpretable? It seems that the only way to know this is to attempt such an interpretation. (In order to verify that a sugar-cube is soluble, we drop it in a glass of water.) But such an attempt is, as before, an intentionally saturated act, i.e., I must be intentional to discover that you are interpretable from the intentional stance and, hence, intentional. Yet how are we to know that *I* am intentional? Doesn't this, the objector inquires, begin the regress again?

Although I share the determined objector's suspicion that something peculiar may be at work in this stage, I am fairly certain that it cannot be for the reasons just advanced. The determined objector, in running an argument such as that sketchily described above, commits an error which amounts to a metaphysics/epistemology confusion. Intentionality, for Dennett, depends entirely upon a claim that something is interpretable in a particular way. In principle, it does not matter at all that nothing in all the universe can discover this. It only matters whether or not that thing could be interpreted if there were such a thing that *could* go about interpreting it. Furthermore,

even if there is such a thing, it is unnecessary and irrelevant if it can and does make such an interpretation. It is an easy slip to make, substituting an epistemological criterion where only a metaphysical criterion is warranted; nevertheless, it remains an error. Consequently, the intentional regress objection remains ineffectual.

(v.) Concluding Remarks

A few concluding remarks are in order to draw to a close this critique of Daniel Dennett's project of offering a theory of intentionality consistent with our contemporary physicalist cosmology. Having reviewed two important problems which beset all attempts to provide such a naturalized account of intentionality, having examined in particular Dennett's own account as presented in *The Intentional Stance* and the way in which he intends to deal with those problems, and having considered one of the more promising criticisms of that account, it is due time for me to provide some response, however vague, to the Dennettian position.

On the whole, the evidence against the intentional stance which I have considered here is marginal at best. Nevertheless, absence of negative evidence is not itself positive evidence. I consider the case in favor of Dennett's intentional stance to be persuasive, often in the extreme, yet I am reluctant to suggest that it is decisive. Moreover, although I have attempted to resist "intuition-mongering" whenever possible in this analysis, many of my intuitions place me solidly opposed to the primary claims which Dennett espouses. Like Stich (1981), I find it more likely to be the case that when we ascribe a belief to some person we are projecting ourselves into what we imagine that individual's state of mind to be than that we are merely positing some usefully fictitious entity which is explanatorily convenient and hence true. Predictive and explanatory success counts just as strongly for the former view as the latter. While mere intuitions do not themselves constitute a case for or against much of anything, they do figure significantly for me in gauging my degree of confidence in a theory. Thus, they shall count, though I have spoken of them little, in my response to Dennett's position.

A Critique of the Intentional Regress Objection

Michael Scriven (1966) defines agnosticism in the narrow sense as a suspension of judgment in which one ought to “make no commitment either way; treat each alternative as approximately equally serious.” This response is appropriate when one possesses substantial support for a position yet has “substantial alternatives still open; a balance of evidence for and against; about 50 per cent probable.” Where I currently stand with regard to the intentional stance, I view just such a field. Dennett’s account is promising, perhaps as promising an account (of those with which I am familiar) as has yet been explicitly developed. But that promise is balanced by the disservice it does to select intuitions and by the availability of similarly promising avenues. With cognitive theory still in its early stages, it seems difficult to do more than narrow the field into plausible and implausible and regard the plausible positions with some measure of agnosticism. In response to Dennett’s theory of intentionality, then, one would do well to *interpret* me as agnostic, for I am so *interpretable*, or I *believe* that that is the case at any rate.

Part Two

CONSCIOUSNESS

“The most central feature of mind, the ‘phenomenon’ that seems more than any other to be quintessentially ‘mental’ and non-physical, is consciousness.”

– Daniel Dennett (1969, p.99)

FENCING WITH JACKSON AND DENNETT (What Mary *Could Not Know*)

Are qualia incompatible with physicalism? Do we obtain information through the phenomenal character of our experiences which is incapturable propositionally? In this paper, I shall examine Frank Jackson's (1982) popularly discussed *Knowledge Argument for Qualia*, which purports to demonstrate that the answer to both of the above questions is yes. One prominent criticism of this argument has been made by Daniel Dennett (1991, pp.398-406). Dennett urges that Jackson's argument is little more than an intuition-pump, and that a proper analysis of it reveals quite the opposite of Jackson's conclusion—the answer to both of the questions is no. Upon review of this criticism, I shall find it wanting. Nevertheless, I shall claim that Jackson's argument is indeed unsound and will provide an account of what I believe Dennett ought to have argued in order to establish this, taking care to consider what might motivate Dennett to prefer his response to mine. I will also consider whether or not his concerns may be accommodated within the criticism which I advance.

(i.) Jackson's Lunge—"The Knowledge Argument for Qualia"

Jackson is a self-proclaimed "qualia freak." Just what sort of characterization constitutes a satisfactory definition of qualia is a much debated issue; nevertheless, there are some indirect approaches to such a definition which play upon the intuitive ease with which we are able to identify them. A quale (the singular of 'qualia') is typically taken to be the phenomenal character of some various sensory input, or the "raw feel" of such an input. We can design machines which identify colors by the wavelengths of various frequencies of light, yet when we see colors there is something that it is *like* to see them (the "raw feels" of seeing blue are quite distinguishable from those of seeing yellow). The color of a ripe, red tomato, the sweet fragrance of a newly-blossomed rose, the taste of a delectable home-cooked meal, the sound of the thunder and beauty

of a symphony, and the feel of the caress of the wind against our skin are all examples of qualia.¹ By “freak” we can understand Jackson to be rather fixated upon them and the importance of accounting for them in whatever conceptual scheme we adopt.

It is evident why he is so fixated upon qualia; Jackson believes that qualia do not fit squarely into a physicalist’s conception of the universe. Yet physicalism is the “going” conceptual scheme. “[T]here are certain features of the bodily sensations especially,” he believes, “but also of certain perceptual experiences, which no amount of purely physical information includes.”² That is to say, given all the information capturable in the physicalist’s terms, we would still be left without an understanding of the relevant features (the qualia) of “the hurtfulness of pains, the itchiness of itches, pangs of jealousy.”³ Thus, physicalism must be false.

This, Jackson believes, is “a perfectly good argument”; he admits, however, that it is weak polemically, as it relies on a premise that is not intuitively obvious to all. Is it possible to state an elaborated version of this argument “whose premises are obvious to all, or at least to as many as possible”?⁴ Jackson believes so. He offers two formulations which putatively meet this stipulation, but really they are the same argument dressed up in different rhetorical guise. The version which seems to be a fan-favorite among philosophical circles is the second formulation, that of Mary the brilliant scientist, so I shall repeat that argument here (with the understanding that the formulation involving Fred’s unique visual discriminatory abilities is actually the very same argument).

Jackson’s argument is this: Mary is a brilliant neuroscientist who has conducted her entire scientific career—indeed, her entire life—in a very controlled environment in which she has only seen objects in black, white, and shades of gray. She has never personally experienced any color aside from these. However, in the course of her meteoric career, Mary has come to acquire “all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky,

¹The feeling of pain and of emotions such as sorrow, happiness, and so forth are also examples of qualia; qualia are not connected only with our traditional five senses.

²p.127

³*ibid.*

⁴p.128

and use terms like ‘red’, ‘blue’, and so on.”⁵ Yet it seems “just obvious,” Jackson claims, that when Mary first leaves her controlled environment and views the great colorful world she will learn something new “about the world and our visual experiences of it.”⁶ But if physicalism is true, then all (correct) information is physical information. And, *ex hypothesi*, Mary already has all the physical information. Thus, if physicalism were true, Mary couldn’t *learn* anything new. But it seems very strongly that she does. Hence, physicalism is false; there is more information than just physical information.

In a condensed and more manageable form, the argument looks like this:

P1: All information is physical information.

P2: Mary knows all the physical information.

But,

P3: Mary does not know all the information about perceptual experience (namely, the qualia associated with such experience).

Therefore,

C1: It is not the case that all information is physical information.

And, since **P1** is central to physicalism,

C2: Physicalism is false.

Jackson’s argument is a *reductio ad absurdum*, intended to demonstrate the falsity of **P1**. **P1** through **P3** thus form an inconsistent set. In order to restore consistency, the proposition contained in one of the premises must be negated. Since **P2** is true *ex hypothesi* and **P3** seems “just obvious,” **P1** (Jackson concludes) must be false, as stated in the conclusion. This is a valid argument. Thus, for the conclusion to be true it need only be the case that premises two and three are true. If Jackson is correct that **P2** and **P3** are indeed true, then clearly it spells trouble for any ambitions of providing a unified physicalist theory, for there is some variety of information (that concerned with qualia) which such a theory is in principle incapable of providing. But is it obviously the case that both of these premises are true? Dennett does not think so.

⁵p.130

⁶*ibid.*

(ii.) *Dennett's Parry—"The Illusion of Imagination"*

Dennett decries the thought experiment upon which Jackson's argument is developed, characterizing it as an "intuition pump that actually encourages us to misunderstand its premises!"⁷ Its point is clear, but the *image* necessary for establishing that point, Dennett believes, is simply erroneous. When we perform the thought experiment, do we actually attempt to imagine that Mary knows *everything*, he asks, or do we merely settle (since the former is just not possible for us to do) for merely supposing that she knows *a lot*? Dennett thinks that the differences imposed upon our imaginations by these two cases are extremely relevant to the outcome. For if we merely imagine that Mary knows a lot, perhaps all that the contemporary sciences have to offer, then it hardly seems surprising that she would learn something new. But, if we are really careful to follow directions, then we shall come to realize that it is perfectly imaginable that when Mary comes out into the world and has a visual experience of the color red for the very first time, she will learn nothing; she already knew it all.

This amounts to a denial of **P3**, the claim that Mary does not know all the physical information about perceptual experience. It would seem that Dennett and Jackson differ with regard to what "seems just obvious." If premise three is false, then there is nothing particularly alarming about the first premise's claim that all information is physical information. All information is physical information, and Mary really does know it all.⁸ But is it altogether clear whether **P3** is true or false; has Dennett so easily freed us from the illusion of imagination? I'm not so certain.

Recall that what is at issue is the claim that all (correct) information is physical information. If there is any information which is not "physical," then physicalism is false. Jackson claims that qualia are going to slip the physicalist's net, for they in fact are/convey information which is not physical in character. Dennett, on the other hand, claims that Jackson is mistaken (and so are we if we believe him); if we know *all* of the things known by a utopian science, then we will know qualia/know the information conveyed by qualia, as well. As Dennett would have it, then, qualia

⁷p.398

⁸A novelty at social gatherings, though not conducive to the forming of enduring friendships.

are perfectly explicable within the domain of physical science. But isn't this the very question at issue? Dennett has accused Jackson of intuition-pumping, and it is true (I will agree) that Jackson is guilty of this charge. But Dennett is no less guilty. Jackson begins with the intuition that physicalism and qualia are incommensurable—that there is no manner in which the propositional explications of the physical sciences could confer the knowledge gained through acquaintance with, or experience of, qualia—and relates a story which relies upon our sharing this intuition to foster consent to its conclusion. Dennett, on the other hand, begins with the intuition that physicalism and qualia are quite commensurable and relates a retelling of Jackson's story which seems to support or vindicate his intuition, but *only* if we share the intuition to start. A brief examination of Dennett's continuation of Mary's misadventures will reveal that this is the case.

Dennett offers us his imagining of the way Mary would react to her first experience of viewing a colored object, a blue banana intended to trick her. Mary looks at the banana and accuses her captors of attempted trickery; bananas are yellow not blue. She then proceeds to explain to the baffled tricksters that she knows everything that could be known about the physical causes and effects of vision. Knowing what effects yellow and blue objects would have on her nervous system, she knew exactly what thoughts she would have when confronted with the appropriately colored objects. Knowing vast amounts of information about her reactive dispositions allowed her to know in the end that the banana before her was blue, not yellow.

Dennett's story is articulate and clever, but is it to the point? I claim that it is not. For consider these three possibilities: Mary sees the blue banana and says, "Oh, so that's what it's like to see yellow" (Jackson's version); she sees the blue banana and exclaims, "Hey, bananas are yellow, but this one is blue!" (Dennett's version); and Mary sees the blue banana and observes, "Well, this obviously is a chromatically-challenged banana, since bananas are yellow and not blue, yet it's nice to finally know what blue looks like" (Dusek's version). Dennett seems to want to claim that *all* it is to acquire the knowledge associated with qualia is to acquire knowledge of reactive dispositions associated with particular sensory experiences. But to allow him that is to allow him to beg the question against Jackson. Jackson claims that there is something beyond all of this which we learn

when we have a sensory experience (there's the "what it's like," recall). Note, then, that the story I tell of Mary's response allows Dennett the legitimate claim that one could, in theory, recognize what type of visual experience one is having by noting reactive dispositions, without begging the question against Jackson by supposing that such recognition exhausts the knowledge acquired by having such a visual experience. It is compatible that Mary could both be aware that she is having a visual experience of the color blue *and* come to learn *what it is like* to have such a visual experience (become associated with the qualia). So Dennett's story does not make it quite as clear as may be thought that one could somehow deduce or infer or imagine the qualia of a visual experience from the physical facts—be they reactive dispositions or something else—related to such an experience. It may be *possible* that it could be done—after all, who can say what may count as physical facts in a utopian science?—but it is far from *obviously* possible, especially on the basis of the story considered above.

We are left, then, with Jackson's claim that qualia are radically different from the things that make it into a physicalist's cosmology and Dennett's claim that the phenomena we associate with qualia are perfectly explicable under, or commensurable with, such a cosmology. The claims are certainly competing, and there is no obvious way to choose between them. This stalemate may appear to be a failing of Jackson's argument; Dennett, it seems, would agree with such an analysis: "My point is not that my way . . . proves that Mary *doesn't* learn anything, but that the usual way of imagining the story doesn't *prove* that she *does*."⁹ But this is surely a much weaker criticism than would be desirable. Jackson's argument is not obviously sound, but then it is not obviously unsound either. There still seems to be room for rationally doubting physicalism *on the basis of* the knowledge argument for qualia.

I conclude that Dennett's strategy of attacking P3 yields somewhat inconclusive results. If the Knowledge Argument for Qualia is to be defeated, it will not be by arguing about what Mary does or does not learn. While Dennett's point is well-taken, I believe that a more cogent criticism may be formulated. With that said, I will turn now to a consideration of what Mary actually *knows*.

⁹p.400

(iii.) *Dusek's Parry—"A Myth of Physical Information"*

There is a temptation to attack Jackson's argument on several counts of lack of clarity or precision. For instance, one might hold grave reservations about the notion of "information" which Jackson invokes and relies upon. Or, similarly, many physicalists would likely object to Jackson's characterization of physicalism. "Jackson's argument defeats *this* brand of physicalism; we'll give you that," they might concede. "But who holds to that version of physicalism, anyway?"¹⁰

Pushing hard on either of these may well cause Jackson's argument to topple. Even overlooking that goal, it is still of great importance to be careful about our treatment of each. Nevertheless, I think that there is an intuitive appeal to the Knowledge Argument that is not necessarily dealt with sufficiently by such a response. Philosophical precision aside, the thought experiment on which the argument is based seems itself to mount a compelling case, at least at the level of gut instinct. For this reason, I propose that we overlook for the time being any misgivings over these potential weak points and grant Jackson his portrayal as approximately correct. I intend to demonstrate that we may still undermine the Knowledge Argument, as advanced and without appealing to dubious opinions on imagination, by challenging the *ex hypothesi* claim of **P2** that Mary knows all the physical information.

What would be required to undermine the truth of this premise? How might we legitimately reject it on charges of falsity, when it is simply taken by hypothesis to be true? Clearly we cannot do so by method of factual disconfirmation. The argument is based upon a thought experiment, one quite obviously impossible to replicate in the real world. But this does not count against the argument. Yet, if we cannot disconfirm it by factual evidence, and we have no reason to suppose that the premise itself is incoherent, what option remains?

¹⁰Lenny Clapp related once that physicalism has become much like existentialism. It's hard to obtain more than a *Zen*-like impression of what either really amounts to.

There is one. And it is not to be overlooked, for I think it the key to freedom from the Knowledge Argument. If we can cast reasonable doubt on the very *possibility* that such a proposition is true, then it is immaterial that the premise is assumed to be true *ex hypothesi*, or that we can only grapple with it in the arena of the thought experiment. Even the gratuitously assumed cannot make the impossible true.

But what reason have we to doubt the very *possibility* of Mary's knowing all the physical information (by the provisions of the thought experiment, remember)?

I propose that we may reject premise two because it is reasonably doubtful that, given the parameters of the thought experiment, Mary could actually have obtained *all* the physical information, brilliant neuroscientist though she is. Mary's captivity itself prevents this. How so?

Jackson informs us that Mary knows all the information of all the physical sciences in a utopian academic world. Clearly, then, she knows all the propositions concerning the physical world and the phenomena and processes within it. Or, Mary has all the *propositional knowledge*. Jackson seems to think that this, for the physicalist, is all it is to know everything. If one has all the propositional knowledge, or propositionally capturable information, then one has all the knowledge there is to have. But this cannot be correct. For what reason is the physicalist committed to this view?

It is perfectly open to the physicalist to maintain, for one thing, that not all features of the physical world are epistemically accessible. We may not be able to know anything about a given object or phenomenon, and as such may not be able to express anything about it, but this does not mean that said object or phenomenon is not physical. This itself does not threaten Jackson's position, for Jackson requires only that the information about those things of which we *are* capable of having knowledge must be physical. If we acquire information, that information must be physical. This, however, does not show that that information must be propositional, which is the conclusion that Jackson needs. Just as the physicalist can toe the line above, she may likewise hold that most of our propositional information may only be acquired/formulated through having the requisite sorts of experiences. (Some hypotheses in particle physics one simply does not

falsify without the aid of a super-collider and “observation” of the resulting collisions.) But what is to prevent her from holding that much of the requisite sorts of experience, and the information so acquired, is simply not the sort of *stuff* that *can* be formulated in the propositionally land-locked domain of the physical sciences? And how does this show that the *phenomenal* or *experiential* sort of information does not meet with the physicalist’s conception of the world, that it is not itself “physical”?

If this is correct, if there is room to suppose that the notion of experiential information is as important to an analysis of Jackson’s thought-experiment and Knowledge Argument as the notion of propositional information, then it is easily recognizable how P2 could be simply impossible. It is open to even the most thoroughgoing physicalist to object to the claim that Mary could have obtained all the physical information while locked in her color-proof room. There is, this physicalist will object, no way in which she could have acquired select types of information about the experiencing of color—the *what it’s like*, particularly—if she only can obtain propositional information. It may be that the only way to know what it is like to view a ripe, red tomato is to view a ripe, red tomato. But this, surely, is no threat to physicalism (or, at least, not the manner of threat Jackson makes it out to be).

Jackson, in the last analysis, has not proven that physicalism is false. He has merely shown that not all physical information is propositional. But there is a great gulf to cross in showing that the latter amounts to the former. I suggest that Jackson has not done so; neither is it clear that he could. If we are going to cooperate with Jackson in using his notion of information, what reason have we to suppose that there are not different varieties of information—*propositional* as well as *experiential*—with sometimes diverging domains of knowledge, yet both of which concern only objects that square with a physicalist’s ontology? Are qualia any more troublesome than propositions or, for that matter, information? Jackson has not made it clear that they are.

APPENDIX A

“Is he willing to prevent evil, but not able? then he is impotent. Is he able, but not willing? then he is malevolent. Is he both able and willing? whence then is evil?”

– Epicurus (in Pojman 1994, p.163)

POWER, BENEVOLENCE, AND SUFFERING: THE MORAL OBLIGATION NOT TO PROCREATE

The problem of evil has long stood as one of the central issues in the debate over the rationality of religious belief, particularly the rationality of theistic belief. The idea is that there is something incoherent in believing certain propositions about God (e.g., that God is omnipotent, omniscient, and omnibenevolent) and also believing that there is evil in the world. When the problem of evil is asserted as an argument against the existence of God, this is known as an argument from evil. Though commonly referred to as *the* argument from evil, such a title can be misleading, for there actually can be, and is, more than one such argument. For example, the Miltonian version of the argument from evil is formulated with the presence of sin as the feature of reality which is difficult to reconcile with the theistic conception of God. How can a good and loving God allow creatures to sin? But that is a very different formulation from the argument from evil popularly discussed in contemporary literature on the subject. This argument is advanced with the presence of suffering as the feature of the world which is difficult to reconcile with the existence of an all-powerful, all-knowing, and perfectly benevolent God. Specifically, it is needless suffering, the sort which does not serve some higher purpose such as the prevention of greater suffering or the bringing about of some better state of affairs, which is the crucial variety of suffering. Terminal cancer of the throat in a young child is a relatively uncontroversial case of such needless suffering. It is this latter formulation, the argument from suffering, that is relevant to my project in this paper.

While the import of atheistic arguments is discussed frequently enough insofar as those arguments pertain to the alleged rationality of religious belief, it seems that their relevance to other matters is sometimes overlooked. In this paper I want to consider the relevance of the argument from evil to a very pervasive aspect of our lives, that of having children. I will argue that we cannot consistently affirm the argument from evil as sound *and* believe it morally permissible to have children. To establish this claim, I shall argue that the crucial premise of the

argument from evil is merely a particular instantiation of a more general principle, which I shall call *The Principle of Benevolent Action*. I shall then demonstrate that this principle may be re-instantiated as a premise which motivates an argument parallel to the argument from evil. I shall call this argument *The Benevolence Argument Against Procreation*. The conclusion of this argument will be that human agents who have children fail to be benevolent. As we ought to be benevolent, the result will follow that if we believe the argument from evil is sound then we ought not to have children.

(i.) *The Suffering Argument for Atheism*

The suffering argument for atheism consists of five premises and a conclusion:

P1: God is an omniscient being.

P2: God is an omnipotent being.

P3: God is an omnibenevolent being.

P4: A being which is conjointly omniscient, omnipotent, and omnibenevolent will not allow people to suffer unnecessarily.

But,

P5: People suffer unnecessarily.

Therefore,

C: Either God is not omniscient, or God is not omnipotent, or God is not omnibenevolent. (The traditional theistic God does not exist.)

Note that this is essentially a *modus tollens* argument.¹

The first three premises of the argument are descriptive in character; we are to understand God to have at least these three properties. This gives us the classical (Judeo-Christian) conception of the theistic God. The fifth premise is a factual premise; it is an observation about the state of affairs in the world. However, we cannot obtain the conclusion from these four premises. The conclusion follows only from a tension between the conjoint attribution to God of

¹If a being with a particular set of properties exists, then people do not needlessly suffer. But it is not the case that people do not needlessly suffer, i.e., people needlessly suffer. Therefore, it is not the case that a being with the stipulated set of properties exists.

the properties in premises one through three and the factual claim in premise five. The fourth premise gives us this tension.

It is open to the critic of the suffering argument for atheism to question the truth of any of its premises. It is doubtful, in theory at least, that people suffer needlessly; perhaps all suffering *is* necessary. One may thus question premise five. Nevertheless, I think the denial of **P5** quite implausible and a desperate act. Moreover, I find nothing philosophically interesting in such a maneuver, so I will not consider it here. Alternately, one might hold any of premises one through three suspect. There is much debate and discussion about the attributes of God, should God exist. Giving up God's omnipotence, omniscience, or omnibenevolence will prevent the *reductio ad absurdum* on which the argument from suffering depends. While there is nothing inherently implausible about denying **P1**, **P2**, or **P3**, I find such a move also to be uninteresting insofar as this paper is concerned. There is, however, a premise which it will prove fruitful to subject to rational scrutiny. In premise four it is asserted that a being with the properties of omniscience, omnipotence, and omnibenevolence will not allow people to unnecessarily suffer. There is much to reap, I believe, from a consideration of this claim.

(ii.) The Principle of Benevolent Action Abstracted

Why would a being which is conjointly omnipotent, omniscient, and omnibenevolent not allow people to suffer needlessly? In considering this question, I shall reveal a general principle, the principle of benevolent action, which relates an agent's power to prevent suffering to the way that agent must act if she is to be benevolent. Premise four, I shall show, is an instantiation of this principle.

By the premise under investigation, there are four claims which taken together form an inconsistent set: that God's power is unrestricted, that God's knowledge is boundless, that God is unfailingly benevolent, and that people suffer unnecessarily. What is it about this conjunction which makes the set inconsistent? As I shall demonstrate, the inconsistency results from the relationship between benevolence, knowledge, and power and how an agent who is benevolent,

relevantly knowledgeable, and relevantly powerful will act with regards to instances of needless suffering.

What do we mean when we attribute the property benevolence to some agent? We mean at least that that person is inclined to act kindly towards others, that in her actions she takes into account the feelings and desires of those whom her actions will affect; in short, we mean that she acts with regard to the potential happiness and suffering of others. Even at this intuitive level, then, the concept of benevolence is intimately connected to the problem of suffering. But this is imprecise, and more can be said about this connection. Through the use of several examples, I will explicate some important features of the notion of benevolence and its relation to an agent's power and knowledge.

Benevolence is Dispositional. Mother Teresa is benevolent; I suspect that this is a fairly uncontroversial claim to make. She is a woman who has devoted her life to alleviating the pains and sufferings of others. In fact, she spends great amounts of time on just that task. But she sleeps, she eats, she bathes. Presumably there are moments when she simply is not actively preventing somebody's suffering. Yet we continue to call her, and think of her as, benevolent. Moreover, there are other people, people who manage professions, families, and personal activities yet still find ways to give themselves to the needs of others, whom we would be inclined to consider benevolent. Certainly they cannot help as many people as, nor can they always help them to the extent that, Mother Teresa does. But we think of them as benevolent, nevertheless. Yet with these people, and even with Mother Teresa herself, there are occasions when they are not actively preventing somebody's suffering. The reason that this is the case and we still consider them benevolent is that benevolence is a dispositional property.² That is, somebody is benevolent if that person *is disposed* to act to relieve suffering in the appropriate circumstances.

²It may be that all properties are dispositional. I will not engage such an issue here. It matters for now only that benevolence is, in fact, dispositional.

Benevolence, then, is similar to the property of solubility. Solubility, too, is a dispositional property. A sugar-cube will dissolve if placed in room-temperature water. A sugar-cube placed on a non-aqueous table, on the other hand, simply sits and most definitely does not dissolve. Nevertheless, a sugar-cube is soluble because it is disposed to dissolve under certain conditions. We would not expect sugar-cubes to always be dissolving, even when they are not in a liquid, in order to deem them soluble. In the same way, a person who takes the time to perform some recreational activity such as reading or biking or running is not *thereby* non-benevolent. She fails to be benevolent if she is not *disposed* to prevent suffering. This, of course, implies that there are certain conditions under which one must act to prevent suffering in order for one to be benevolent. The next several examples clarify such conditions.

Conditions of Benevolence. Consider the following scenario. A flaming building is rapidly being consumed in a raging inferno. Specifically, this building is a private residence, and several small children are trapped within. If unassisted, they will surely suffer agonizing deaths. Is it possible that I may fail to prevent their (unnecessary) suffering and yet be benevolent? Of course it is.

Ignorant Benevolence. For one thing, I simply may not have any knowledge that this horrible thing is occurring. If I do not even know that there is a burning building somewhere with small children trapped within, then I can still be benevolent, even if I do not act to save the children. A terribly large number of people suffer from many different tragic occurrences. I know of only a minuscule fraction of these cases, at best. We would not say, I hope, that I am not benevolent on account of my failure to prevent the suffering that I did not even know about. It is how I act with respect to the instances which I *do* know about that is relevant to my benevolence. Thus, knowledge is a crucial element of the power condition of benevolence.

Impotent Benevolence. Alternately, I may lack the appropriate ability to relieve another's suffering. I may actually walk by that very building and hear the cries of the children within. But all the entrances to the building have either collapsed or have been engulfed in impenetrable flame. The fire department is on its way, yet it cannot possibly arrive in time. There is not a

thing I can do. I might stand in paralyzed horror, traumatized by my own inability to spare these children their agony. I am doing nothing to prevent the suffering of which I have very good knowledge. Still, we cannot say that I am not benevolent on account of this, for I have no options available to me on which to act. I am, so to speak, impotent with respect to the suffering at hand. It need not even be the case that I stand idly. Perhaps I break down a flaming door and charge into the furnace but am overcome by the heat and smoke within moments and pass out. Once again, through a sort of impotence, I am unable to act so as to prevent the children's suffering. Relevant power, then, is another key condition of benevolence.

Knowledgeable, Potent Non-Benevolence. There is one more important case to consider. I pass by the same flaming building. I hear the children's cries. Upon surveying the scene, I find that I may gain entrance through a back door. In one of those bizarre, chance events, I just happen to be wearing my best asbestos suit. Can I pass by without failing to be benevolent? It no longer seems that I can do so. There is an obvious case of suffering before me. I have the knowledge of its occurrence, and it is within my ability to prevent it. If I fail to act so as to prevent their suffering when it is within my power to do so (I have the knowledge and the ability), then can I still be benevolent? No, it is clear that I cannot. The only way I can be benevolent in this situation is by taking action to prevent the children's suffering.

Thus, knowledge and power relevant to the prevention of a particular instance of suffering are necessary conditions of benevolence.³ In other words, when we say that someone is benevolent, we mean that she is disposed to act in a certain way when the appropriate conditions obtain. Those conditions are that she has the relevant knowledge of an instance of needless suffering and the power relevant to preventing that suffering. When both of these conditions obtain, she will act to prevent the suffering at hand. We now have, from the examples and analysis offered above, all which we need in order to state the principle of benevolent action, abstracted from its instantiation as premise four of the suffering argument for atheism. This principle is:

³Note, however, that I am not claiming they are sufficient conditions as well.

PBA: If x is benevolent, and x has the relevant knowledge, and x has the relevant power, then x will act so as to prevent unnecessary suffering.

If it is indeed the case, as I have claimed, that premise four of the suffering argument for atheism is an instantiation of the principle of benevolent action, then we ought to be able to demonstrate this by performing just such an instantiation. And, in fact, we can. What happens when we take God as the agent x in the principle of benevolent action? **PBA** tells us that if God is benevolent, and God has the relevant knowledge, and God has the relevant power, then God will act so as to prevent the suffering to which this knowledge and power is relevant. Since God by **P3** is omnibenevolent, or infinitely/perfectly benevolent, God will always be benevolent; God will always prevent suffering that God knows about and has the power to prevent. But by **P1** God is omniscient; God knows of all suffering (with complete detail). And by **P2** God is omnipotent; that is, God has the power to prevent all unnecessary suffering. Thus we arrive at the conclusion that God, a being which is conjointly omniscient, omnipotent, and omnibenevolent, would not allow people to needlessly suffer (for God would always prevent such suffering). The principle of benevolent action re-instantiated yields the conclusion exactly as desired, an indication that my abstraction has been successful.

(iii.) The Benevolence Argument Against Procreation

The principle of benevolent action is general. By that I mean to say that it is not a principle which is specific to beings like God, or is applicable only to arguments against the existence of particular divine beings. Its scope is (and needs to be) broad enough to include all agents to which we may meaningfully attribute the property of benevolence. Benevolence is a virtue applied to human beings, primarily, so the principle of benevolent action needs to be sensible in the context of human affairs.

At the outset of this paper, I promised to argue in support of the claim that people ought not to have children, and I am now prepared to make good upon that promise. I indicated that a

general principle, which relates an agent's knowledge and power to prevent needless suffering to the way in which she must act in order for her to be benevolent, is abstractable from the suffering argument for atheism in which it is instantiated. I have established that this is true. I further indicated that this principle may be re-instantiated as a premise in an argument parallel to the suffering argument for atheism. I shall do this now.

Nobody is born who does not suffer. The claim that any person born into this world will suffer is undeniable. So, the act of procreation is an act which ensures that suffering will occur. As potential parents, then, in considering whether or not to have children, we are making decisions (explicit or implicit) concerning a very important issue; we are, in effect, deciding whether or not to place a person in an environment in which that person's suffering is guaranteed (it is only a matter of degree). But we know that we have the ability not to do so in virtue of the very fact that we consider *whether* to have children. Human procreation is an optional endeavor; it is preventable. And because procreation is preventable, we have the power to prevent the needless suffering of those who would be born, by not having children. (At the very least, we have the *ability* to prevent the suffering. But do we necessarily have the *knowledge* that such suffering will occur? Perhaps we have not entertained the possibility of such suffering and as a result are open to the claim of ignorance. Well, we are ignorant no longer. Upon hearing or reading this paper, the issue has been considered and the relevant knowledge obtained.)

I advance here, then, the benevolence argument against procreation:

P1: I have relevant knowledge (that my children would suffer if they were born).

P2: I have relevant power (to prevent the suffering of my potential children, by not having children).

P3: I am benevolent.

P4: If I am benevolent, then I will act so as to prevent the suffering of which I have knowledge when it is within my power to do so.

P5: My children would suffer if they were born.

Therefore,

P6: If I have children, then either I do not have the relevant knowledge, or I do not have the relevant power, or I am not benevolent.

But, **P1** and **P2** are true. So,

C: If I have children then I am not benevolent.

Hence, from the argument above, if one is to be benevolent then one cannot have children.

Moreover, since one *ought* to be benevolent, one ought not to have children. As the suffering argument for atheism and the benevolence argument against procreation are parallel arguments advanced on the basis of the principle of benevolent action, it is inconsistent to affirm the argument from evil as sound and to believe it morally permissible to have children.

(iv.) Criticisms and Their Responses

C: *The suffering argument for atheism and the benevolence argument against procreation are dis-analogous. It is distinctly different to be an omniscient, omnipotent, omnibenevolent agent than it is to be an agent of limited knowledge, power, and benevolence.*

R: The argument advanced in this paper is not an argument from analogy, hence the criticism of dis-analogy is neither here nor there. The suffering argument for atheism is motivated by an abstractable principle of benevolent action. That principle is re-instantiable for any agent(s) regardless of the degree of knowledge or power of that agent. It is the *relationship* between an agent's knowledge and power and instances of needless suffering, not the amount of knowledge or power, that is relevant to the success or failure of arguments such as those advanced here. The benevolence argument against procreation is perfectly assertable without any reference at all to the argument from evil. The presence of the latter argument in this paper merely serves dialectical and polemical purposes.

C: *The formulation of the principle of benevolent action is too strong. It requires that we prevent every instance of needless suffering of which we have "relevant" knowledge and power. But how does this differ from omnibenevolence? The principle of benevolent action is, properly, a principle of omnibenevolent action. But this, surely, demands too much. Mother Teresa, one can reasonably assume, has at some point simply passed by a suffering individual whom she could have aided. Nevertheless, if anyone is benevolent, she is.*

R: I am of the opinion that there are two notions of benevolence relevant to a consideration of this criticism. The first is captured by my principle of benevolent action; it is a descriptive notion concerned only with particular instances of suffering and responses to them. The second notion, on the other hand, is an evaluative notion. Rather than speaking of benevolence with respect to particular circumstances, we generalize over a person's life

(“Mother Teresa is a benevolent person” rather than “benevolence demands such and such an action”). It is the latter notion which the criticism above appeals to. But it is important to note that it depends upon the former notion; we consider how a person “measures up” to the descriptive notion over the course of her life in order to decide that she is not really a benevolent person, or a generally benevolent person, or a very benevolent person. In fact, if a person flawlessly adhered to the descriptive notion of benevolence, then we would be well within our rights in saying that she was omnibenevolent. So I maintain that my principle of benevolent action is properly formulated in terms of the descriptive notion of benevolence, while it does not commit us to denying the benevolence (in the evaluative sense) of a person who is imperfect (that is, any person). But that does not let us off the hook on the procreation issue. The moral obligation is an obligation to what I have called benevolence in the descriptive sense.

- C:** *The principle of benevolent action, though stating necessary conditions of benevolence, does not state sufficient conditions of benevolence. Although this was duly noted in the development of the argument, the neglect of these other crucial elements of benevolence results in a skewed, or at the least oversimplified, account of how we must act in order to be benevolent. For instance, intuitively it seems that benevolence has just as much to do with allowing or propagating instances of joy or happiness as it does with preventing instances of suffering. It seems, moreover, that denying your potential child the joy and happiness she would experience by living is really not warranted by the suffering you will spare her. It is just as un-benevolent to deny your potential child joy as it is to allow her to suffer needlessly.*
- R:** The principle of benevolent action is only formulated with regard to prevention of suffering, and its conditions are explicitly stated to be necessary, though not necessarily sufficient. It is arguable, I will grant, that the complementary conditions alluded to above are often of importance where issues of benevolence are concerned. Although it is possible that the latter conditions (joy conditions) merely collapse into the former (suffering conditions), I will allow for now, for the sake of argument, that that is not the case. Nevertheless, I argue that for simplicity’s sake I am warranted in neglecting the joy conditions, as they have no bearing whatever on the arguments involved in this paper. I argue this for the reason that I believe it is absurd to suppose that I can be un-benevolent towards someone who does not exist, a potential person. Since in the context of the arguments in this paper we are only dealing with the benevolence of procreating, and hence we are only dealing with “potential” people, for simplicity’s sake I opt to streamline the principle of benevolent action and ignore the joy conditions.

- C:** *Given the response above, there seems to be an inconsistency in your treatment of potential persons as opposed to actual persons. The benevolence argument against procreation is formulated with deference to the suffering of your potential children. You are trying to act benevolently, it would seem, towards an individual that does not exist; yet above you decry the possibility that you act un-benevolently towards your potential children if you do not have children (due to your denying them the ability to experience joy/happiness), since they do not exist. What could possibly be the difference between these two cases that warrants your differing treatment of each?*
- R:** What I am claiming is basically this: If I have a child, there will be an ACTUAL person who will suffer (unnecessarily), and I am in a position to prevent that suffering now by not having that child (even though there is not *currently* an object of my benevolence). Yet it makes no sense to talk about what I am denying to this POTENTIAL person, since there is no such person and there *never will be*. What I am claiming, then, is that in the first case my actions result in the suffering of an ACTUAL person, while in the second case nobody suffers at all. In the first case there is an object of my benevolence while in the second case there is none. Benevolence only matters in the context of ACTUAL persons. Moreover, I mistrust the talk of actual/potential persons, given the misunderstanding that it seems to invite. If I urge concern for both people and potential people, I do not mean to suggest that I care about people “here” as well as people “out there” ((hand-waving)) somewhere else. I am merely saying that I care about the people that I live with already and people that I will live with some day. ‘Potential person’ is just intended to be a concise way of saying, “You know, that person, whoever it WOULD be and whatever it WOULD be like, who WOULD breathe and eat and sleep and drink crazy KOOL-AID mixes and watch *The Tick* (but would also suffer unnecessarily)—yeah, that one.” Or, more elegantly, I think that ‘potential person’ has *intension* but no *extension*.

Now, it won’t do to attempt to reformulate my claim above to read: “If you have a child, there will be an ACTUAL person who will experience joy, and you are in a position to prevent that joy now by not having that child (even though there is not *currently* an object of your un-benevolence).” Such an attempt misunderstands my claims, as I will attempt to demonstrate.

We have benevolence. Benevolence means preventing suffering when one can and (granting the objection above for now) NOT preventing joy whenever one can. But, we are only concerned with preventing the suffering and allowing the joy of actual persons, as opposed to some half-baked notion of a person floating around in a quantum-esque indeterminate state.

So the question is, will my opting to have or not to have children violate either of these conditions of benevolence. With regard to having children, I answer that the first condition is violated. With regard to not having children, I answer that *neither* condition is violated.

I see it like this: I'm contemplating having a child. There is no person who could be the object of my benevolence during this contemplation period. So there is nothing inherently important about the way matters stand here. It begins to matter when I consider the *consequences* of either my having a child or not having a child. What are they? Well, if I have a child, then there is an honest-to-God actual person, and that person will suffer unnecessarily. Since I'm able to prevent this suffering, if I don't, then condition one is violated. Now, what if I don't have a child? Well, it might seem that I'm depriving someone of joy (since there could be someone who would experience joy, but I'm preventing that state of affairs) and thus violating condition two. But this is not right, because I'm simply not depriving an actual person of anything at all. Saying that there could be an actual person who would experience joy is just the same as talking about a potential person. Now, this is different from the suffering case, because I'm saying that there *will be* an actual person who *will* suffer needlessly. But with the joy case, one cannot say that there will be an actual person who will be deprived of something valuable if I don't have a child, because by not having that child I'm making it manifestly absurd to suggest that there will be someone who is deprived of joy. At most one can argue that there *would have been* someone who *would have* experienced joy, which is just to say that there *is* no actual person (ever) that we're dealing with at all.

A variety of other criticisms have already been articulated in response to the argument developed in this paper. Many fail on account of their reluctance to address the argument directly. It is a simple matter to mount intuitively compelling reasons that a flaw must be present in my argument, through the use of appeals to the personal fulfillment felt by those who have children or to seemingly undesirable consequences which result from refusal to procreate (the end of the species, for example). But such criticisms fail to appreciate the plausibility of each premise on the basis of which the argument is advanced. If my conclusion is false, the argument itself merits more than casual dismissal nevertheless. For it is certainly valid, and if the conclusion is false it still seems far from obvious which premise must likewise be false. And, of course, the falsity of a premise in the argument which I advance may pose serious implications

for the argument from evil, an argument widely found compelling in analytic circles. In the final analysis, however, I believe the strength and value of this argument lie not in its conclusion but in the Scylla of philosophical investigation and the Charybdis of practical intuition which it compels us to navigate: we can have what we want or we can have what we need, but it would seem that we cannot always have both.

APPENDIX B

“To be sure, thought does not always reflect reality accurately, but it is the function of logic always to work toward a more adequate relation between language and reality.”

– S.E. Stumpf (1993, p.84)

BIBLIOGRAPHY

- Armstrong, David M. 1970. "The Nature of Mind," in Hoy and Oaklander, eds., *Metaphysics* (Wadsworth, 1991). pp.234-40.
- Baker, Lynne Rudder. 1994. "Content Meets Consciousness." *Philosophical Topics*, 22, pp.1-22.
- Best, John B. 1995. *Cognitive Psychology*. Minneapolis/St. Paul: West Publishing Company.
- BonJour, Laurence. 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- _____. 1991. "Is Thought a Symbolic Process?" *Synthese*, 89, pp.331-52.
- Churchland, Paul M. 1985. "Reduction, Qualia, and the Direct Introspection of Brain States." *The Journal of Philosophy*, 82, pp.8-28.
- Churchland, Paul M. and Patricia S. 1990. "Could a Machine Think?" in Hoy and Oaklander, eds., *Metaphysics* (Wadsworth, 1991). pp.286-95.
- Dennett, Daniel C. 1969. *Content and Consciousness*. London: Routledge & Kegan Paul.
- _____. 1971. "Intentional Systems." *Journal of Philosophy*, 68, pp.87-106.
- _____. 1978. *Brainstorms*. Cambridge, MA: MIT Press.
- _____. 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- _____. 1987. *The Intentional Stance*. Cambridge, MA: MIT Press.
- _____. 1991. *Consciousness Explained*. Boston: Little, Brown.
- _____. 1994. "Get Real." *Philosophical Topics*, 22, pp.505-68.
- _____. 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.
- _____. 1996. *Kinds of Minds*. New York: Basic Books.
- Gardner, M. 1970. "Mathematical Games." *Scientific American*, 223, no.4, pp.120-23.
- Herbert, Nick. 1993. *Elemental Mind*. New York: Plume.
- Hick, John. 1977. "Evil and Soul-Making," in Pojman, ed., *Philosophy of Religion* (Wadsworth, 1994). pp.177-81.
- Hofstadter, Douglas R., and Dennett, Daniel C. 1981. *The Mind's I*. New York: Basic Books.
- Jackson, Frank. 1982. "Epiphenomenal Qualia." *The Philosophical Quarterly*, 32, pp.127-36.
- _____. 1986. "What Mary Didn't Know." *The Journal of Philosophy*, 83, pp.291-95.

Bibliography

- Kim, Jaegwon. 1996. *Philosophy of Mind*. Boulder, CO: Westview Press.
- Lormand, Eric. 1994. "Qualia! (Now Showing at a Theater Near You)" *Philosophical Topics*, 22, pp.127-56.
- Mackie, J.L. 1955. "Evil and Omnipotence," in Pojman, ed., *Philosophy of Religion* (Wadsworth, 1994). pp.185-92.
- Martinich, A.P., ed. 1990. *The Philosophy of Language*. Oxford: Oxford University Press.
- Millikan, Ruth. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.
- Nagel, Thomas. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.
- _____. 1986. *The View From Nowhere*. Oxford: Oxford University Press.
- _____. 1987. *What Does It All Mean?* Oxford: Oxford University Press.
- Plantinga, Alvin. 1974. "The Free Will Defense," in Pojman, ed., *Philosophy of Religion* (Wadsworth, 1994). pp.193-211.
- Pojman, Louis P., ed. 1994. *Philosophy of Religion*. Belmont, CA: Wadsworth.
- Putnam, Hilary. 1973. "Meaning and Reference." *The Journal of Philosophy*, 70, pp.699-711.
- Rowe, William L. 1979. "The Problem of Evil and Some Varieties of Atheism," in Geivett and Sweetman, eds., *Contemporary Perspectives on Religious Epistemology* (Oxford University Press, 1992). pp.33-42.
- Scriven, Michael. 1966. "The Presumption of Atheism," in Pojman, ed., *Philosophy of Religion* (Wadsworth, 1994). pp.403-11.
- Searle, John R. 1990. "Is the Brain's Mind a Computer Program?" in Hoy and Oaklander, eds., *Metaphysics* (Wadsworth, 1991). pp.276-85.
- Shepard, Roger M. 1978. "The Mental Image," in Honeck, ed., *Introductory Readings in Cognitive Psychology* (DPG, 1994). pp.83-97.
- Siewert, Charles. 1993. "What Dennett Can't Imagine and Why." *Inquiry*, 36, pp.93-112.
- Stich, S. 1981. "Dennett on Intentional Systems." *Philosophical Topics*, 12, pp.38-62.
- Stumpf, S.E. 1993. *Socrates to Sartre*. New York: McGraw-Hill, Inc.
- Thagard, P. 1996. *Mind: Introduction to Cognitive Science*. Cambridge, MA: MIT Press.
- Turing, A.M. 1950. "Computing Machinery and Intelligence," in Honeck, ed., *Introductory Readings in Cognitive Psychology* (DPG, 1994). pp.15-24.

Bibliography

- van Gulick, Robert. 1993. "Understanding the Phenomenal Mind: Are We All Just Armadillos?" in Davies and Humphries, eds., *Consciousness* (Blackwell). pp.137-54.
- Watt, Stuart. 1996. "Naive Psychology and the Inverted Turing Test." *Psychology*, May 12 (Online).
- Zeidenberg, Matthew. 1987. "Modeling the Brain," in Honeck, ed., *Introductory Readings in Cognitive Psychology* (DPG, 1994). pp.33-42.

Unpublished Resources:

- Colter, L.W. "How Can Bodies Think?"
- _____. "The Causal Closure of the Physical: A Principle Defrocked."
- _____. "Physicalist Theories of Mind: of their variety and their failure."
- unedited, 1996. *The IWINK Proceedings*.

MINDING THE MENTAL

(Abstract)

In the philosophy of mind, there are two significant areas which provoke a great deal of debate and discussion: consciousness and intentionality. Daniel Dennett, a prolific philosopher and cognitive theorist, has staked out controversial positions on each of these topics. It is the purpose of this research project to examine Dennett's theory of consciousness and his theory of intentionality as well as some significant criticisms of each, and to articulate a response to the material examined. The project will lead as well into a consideration of some relevant issues in the philosophy of language and work by Thomas Nagel on problems of subjectivity and objectivity. Additionally, appended to the project proper will be an unrelated argument for the moral obligation not to have children, which will be developed during the duration of the research honors project.

MATTHEW T. DUSEK is best known in the academic setting by his petitions for "temporal latitude" with respect to written obligations and his formulation of inane but compelling arguments—which he often convincingly claims to believe. Dusek first won critical acclaim with his breakout philosophical bestseller "On Tacit Belief," a "major epistemological achievement" in the (implicit) words of *The Lannon Report's* celebrated feature writer Kevin Lannon, and currently the subject of negotiations for lucrative cinematic release on the big screen. *Minding the Mental* is the first in an open-ended series of "epic philosophy." Look for the smash sequel *Reminding the Mental (Why Did I Get A Degree in Philosophy, Anyway?)* at defunct booksellers everywhere in the Spring of '98! In addition to his regular online columns and performances in slapstick philosophy, Dusek is a creative consultant for *IUC Digital Studios* and Best Boy Electric of *The Lannon Report*, as well as a writer, poet, musician, and grumpy owner of an '85 Mustang—which is itself more than mildly grumpy.