

THE DUALIST

Stanford's Undergraduate Journal of Philosophy

VOLUME XIII

WINTER 2007

This issue is dedicated to Alexei, Angela, and Audrey, for their patient help.

THE DUALIST

Volume XIII ■ Winter 2007

Department of Philosophy
Stanford University

Editors-in-Chief

Collin Anthony, Theresa Berens, Marilie Coetsee, Lang Liu, Govind Persad

Assistant Editors

Luukas Ilves, Max Kleiman-Weiner, Hugh Gorman,

Editorial Staff

Collin Anthony	Luukas Ilves
Jill Barber	Max Kleiman-Weiner
Theresa Berens	Lang Liu
Alayna Buckner	Conor Mayo-Wilson
Marilie Coetsee	Govind Persad
Emily Fletcher	Elizabeth Phillips
Hugh Gorman	Michael Pressman
Emily Gustafson	Eve Rips
George Hill	

Graduate Student Advisors

Alexei Angelides, Angela Potochnik and Audrey Yap

Graduate Student and Faculty Reviewers

Alexei Angelides	Alistair Isaac
Shawn Burns	Angela Potochnik
Katherine Dunlop	Shawn Standefer
Ben Escoto	Allen Wood
Daniel Halliday	Audrey Yap

Authorization is granted to photocopy for personal or internal use or for free distribution. Inquiries regarding all types of reproduction, subscriptions, and advertising space can be addressed to: The Dualist, Department of Philosophy, Stanford University, Stanford, CA 94305, USA.

TABLE OF CONTENTS

A Model of Motorsensory Coordination in Enactive Perception

Anatoly Nichvoloda

William Patterson University

1

From Logical to Epistemic Circularity: The Cartesian Circle Remains

Tan Pei-En

National University of Singapore

17

The Ontological Status of Dreams in Berkeleian Metaphysics

Kenny Pearce

University of Pennsylvania

36

Scanlon's Aggregation Problem

Tim Campbell

University of Buffalo

60

An Interview with Michael Walzer

85

Undergraduate Resources

95

Acknowledgements

99

About *The Dualist*

101

A MODEL OF MOTORSENSORY COORDINATION IN ENACTIVE PERCEPTION

ANATOLY NICHVOLODA

James Patterson University

Abstract:

The general idea of perception is the recognition and interpretation of sensory stimuli based on memory and neurological processes that underlie this ability. In this paper I will review and contrast two major approaches to understanding perception: the AI inspired view of perception as passive computational processing of environmental sensory input and the thesis of the embodied enactive perception that considers perception to be proactive motorsensory interaction with environment. Building on the ideas of the enactive perception, I attempt to explicate visual perception as a model of motorsensory coordination. By specifying the falsifiability conditions of the model I seek to demonstrate the empirical nature of the thesis of enactive perception. Reviewed experiments conducted on motorsensory robotic platforms strongly suggest that perception is intimately related to the cognitive agent's ability to initiate and control motor action. Similarly, experiments with human Tactile Vision Substitution Systems provide an empirical demonstration of the necessity to engage in epistemic actions before acquiring ability to perform goal-oriented pragmatic actions. I will argue that the dynamic engagement of a cognitive agent in these two processes constitutes the agent's ability to perceive objects in the environment.

Introduction

Reviewed experiments conducted on robotic platforms strongly suggest that perception is intimately related to the cognitive agent's ability to initiate and control motor action. Similarly, experiments with Tactile Vision Substitution Systems (TVSS) provide an empirical demonstration of the necessity to engage in epistemic actions before acquiring the ability to perform goal-oriented pragmatic actions. I will argue that the dynamic engagement of a cognitive agent in these two processes constitutes the agent's ability to perceive objects in the environment. The purpose of the model of enactive perception as motorsensory coordination that I outline is to provide a useful structure for empirical analysis of complex epistemic and pragmatic activities in terms of the multiple contents and vehicles of the sensory modalities involved in any particular act of perception.

Embodied Perception as an Alternative to Passive Computation:

Embodied perception is the proposal that perceptual systems conceptualize the real three-dimensional world as patterns of possible bodily interactions. It challenges the AI inspired view of perceptual systems as passive stimuli-processing devices running pattern recognition software on modules of hardware. An example of the traditional approach is the Representational Framework for Vision (Marr, 1982). It outlines a system that consists of distinct layers of computational processing that build up complex images in the brain by applying certain computational algorithms to retinal inputs. Marr's theory of visual perception is based on the following assumptions:

- Human vision is a perceptual module with well defined computational boundaries separating it from other perceptual systems and from the higher level thinking systems until processing is complete. That is, the vision module deals strictly with constructing representations of the visual stimuli and makes them available for further processing in higher level cognition;
- The visual perception is in essence representational, which means it is a formal system that can be exhaustively described by mathematical equations;
- The processing of visual signals is conceived of as a sequence of discrete stages, which apply progressively more complex algorithms to yield progressively more complex representations.

According to Marr, a human visual system is a distinct module in the brain that works by constructing a full 3D model of the environment in our heads. From the distribution of light intensities in the retina, cortical edge detectors construct a raw primal sketch in early visual areas such as V1. Further, through recursive algorithms the full primal sketch is constructed by finding complex sets of edges that form contours and boundaries. Later processing adds such features like surface texture, reflectance, and color. Finally, upon completion of all the steps one has a full-blown and detailed 3D view of the surrounding environment.

Instead of assuming that vision consists in the creation of complete internal representations of the outside world whose activation somehow

generates visual experience, the enactive approach to perception put forth by O'Regan and Noe (2001) proposes treating vision as inherently exploratory activity. This approach is based on the idea that vision is a mode of environmental exploration that is implemented with a knowledge of certain visual-motor contingencies. Thus, according to O'Regan and Noe, visual perception is determined by two important aspects:

- There is a set of lawful relations of dependence between visual stimulation and the motor actions of an organism¹;
- Objects, when explored visually, present themselves to us as provoking visual sensorimotor contingencies, corresponding to visual attributes such as color, shape, texture, size, hidden and visible parts.

Therefore, vision is a mode of skillful encounter with the environment that requires implicit knowledge of the sensorimotor contingencies and the ability to make use of that knowledge for the purpose of guiding action, informing thought, and language use. Furthermore, important features of the active visual perception include:

- Cognitive agents acquire the ability to exercise mastery of vision-related rules of the sensorimotor contingencies;
- Knowledge of the sensorimotor contingencies is a practical, not propositional form of knowledge;
- Sensorimotor mastery is context-specific; that is, among all previously memorized action recipes that allow one to make lawful changes in the sensory stimulation, only some are applicable at any given moment.

Thus, the set of action recipes applicable now as you are looking at

1 The difference between visual stimulation and perception can be illustrated by vision restoration cases reported by Sacks (1995). One of his patients, Virgil, upon having his bandages taken off reported light, color and movement all mixed up in a meaningless blur. This would be an instance of visual stimulation vs. Virgil's later acquired ability of visual perception as a result of active interaction with the environment.

this text and the action recipes currently being exercised constitute the fact of your visual perception of this text. It is important to note that the enactive approach to perception does not claim that there are no representations in vision, but rather that their role in theories of perception needs to be reconsidered. According to Noe (2004), it is a mistake to suppose that vision is just a process that employs the computational task-level characterization of retinal input to generate a complete internal model of the world.

Another view that emphasizes the enactive aspect of perception is Clark's (1997) Model of the Opportunistic Mind. It outlines an alternative methodology to the traditional cognitive scientific view of perception as passive computation. He argues that the practice of casting each problem in terms of an abstract input-output mapping and seeking an optimal solution can mislead us in several crucial ways:

- The replacement of real physical quantities with symbolic terms can obscure opportunistic strategies that involve exploiting the real world as an aid to problem solving;
- Conceptualizing the problem in terms of an input-output mapping views cognition as inherently passive computation;
- Search for optimal computational solutions may further mislead us by obscuring the role of the cognitive agent's history in constraining the space of biologically plausible solutions.

So, instead of the old abstract problem solving, Clark offers a new methodology for studying the embodied active cognition, whose key features are:

- Physical quantities with the real-world and real-time focus as inputs and actions as outputs;
- Decentralized solutions where coordinated intelligent action does not necessarily require central planning;
- Extended view of cognition and computation – computational processing is spread out in space and time, often outside the physical body and can incorporate features of the environment.

The basic idea of enactive perception and cognition has found its adherents among many researchers in different fields of scientific inquiry: from Merleau-Ponty's (1945) "vision is palpation with the look" to Brooks' (1991) rejection of the idea of representation as "the wrong unit of abstraction;" from Gibson's (1979) notion of action-oriented affordances to Varela's et. al (1991) focus on repeated sensorimotor interactions between an agent and the world as the basic locus of scientific and explanatory interest. Clark neatly summarizes versions of the idea of enactive perception and cognition in his thesis of radical embodied cognition: "Structured, symbolic, representational, and computational views of cognition are mistaken. Embodied cognition is best studied by means of noncomputational and nonrepresentational ideas and explanatory schemes..." (Clark 1997, 148)

Tactile Vision Substitution Systems: Acquisition of Quasi-Visual Perception with Epistemic Actions

Since the 1960's Paul Bach-y-Rita (1972) has been conducting research devoted to TVSS (Tactile Vision Sensory Substitution). The systems used by Bach-y-Rita in his research consisted of a video camera, portable computer and a matrix of 400 activators: 20 rows and 20 columns of vibrating solenoids. The image captured by the camera is transformed into basic patterns by the computer, and those patterns are represented on the matrix placed on the subject's chest, back, or forehead.

The notion of sensory substitution in a general case denotes the ability of the human central nervous system to acquire the ability to perceive one's environment and successfully interact with it by integrating devices of this sort into the general field of consciousness. In essence, people who acquire such abilities learn a completely new mode of perception. Blind or blindfolded subjects equipped with the TVSS are almost immediately able to perceive simple patterns such as horizontal and vertical lines, detect the direction of movement of mobile targets and orient themselves. While the subject initially feels only successions of stimulations on the skin, after a brief learning process the subject ends up neglecting these tactile sensations, and is aware only of the stable objects at a distance in front of him. In other words, the capacity to recognize basic shapes is accompanied by a mental projection of the objects which are perceived as existing in external space and having certain properties

relative to the subject.

The externalization and intentionality phenomena were later confirmed by a number of experimental observations. For example, when the zoom of the camera is manipulated unbeknownst to the subject, causing a sudden expansion of the representation of the tactile image, “the subject would take evasive action characteristic for people who perceive dangerously approaching objects: they move backwards and raise their arms to shield themselves” (Bach-y-Rita, 1972). Furthermore, visually blind persons discover perceptive concepts that are quite new for them, such as parallax, shadows, and the interposition of objects as well as reproduction of certain classical optical illusions.

The main point of interest for philosophers in this and similar kinds of research is the fact that the subject’s self-initiated and controlled epistemic actions always play a central role in the acquisition of perception that later enables pragmatic actions². Indeed, when TVSS research subjects did not have complete motor control of their cameras (movements up and down, from left to right, zoom forward and back, focusing, etc.) their performance was never better than chance, even after dozens of trials. In other words, perception did not occur when these two conditions were present: (a) the subjects were not able to have full control in manipulating their camera, and (b) if they were not given immediate proprioceptive feedback on their performance in pattern recognition.

Epistemic Actions of Enactive Perception vs. Pragmatic Actions of Behaviorism

The TVSS experiments can be used as a suitable example to ward off a charge of behaviorism that was brought against the thesis of enactive perception by Block (2001). Behaviorism in its general form is the idea that two systems are mentally the same if they have identical input-output capacities and dispositions. Thus, from a behaviorist point of view, a vision restoration subject upon opening her eyes for the first time

² Kirsh & Maglio (1994) define epistemic actions as external physical actions that an agent performs to change her own computational relation to the environment with the purpose of making mental computations easier, faster or more accurate. Pragmatic actions are those that create physical states which physically advance one towards certain goal(s).

should be disposed or able to produce the pragmatic behavior of sitting down (among many others) when exposed to a chair. But, as vision restoration cases observed by Sacks (1995) have demonstrated, until subjects engaged in epistemic actions and acquired the sensorimotor contingencies of a chair, pragmatic action was impossible. Thus, for environmental inputs to acquire their causal power of informing pragmatic action, certain sets of motor-system-caused invariant changes of the environmental stimuli have to be established. From our everyday experience we know that when faced with confusing inputs and unable to exercise pragmatic actions, we revert to epistemic actions.

Thus, epistemic actions, which are at the core of the concept of enactive perception, essentially enable pragmatic actions, and not recognizing this connection is equivalent to leaving the proverbial cart without the horse. Hurley (1998) argues a similar point by saying that behaviorism makes a mistake by ignoring feedback from output to input. In other words, perception and action are related by the dynamic patterns of the reciprocal output-informing-input-informing-output loops. Furthermore, behaviorists have stressed that the cognitive responses of a creature are always under the control of some environmental stimulus property (Skinner, 1957), thus effectively excluding a cognitive agent from the explanatory scheme. In contrast, enactive perception puts an individual in the focus of the investigation by emphasizing the role of the individual's epistemic motor actions in perception and cognition.

Based on the importance of the epistemic motor actions in the acquisition of perception, I suggest that the direction of causation in enactive perception (output-informing-input) should be reflected in the language discussing it. Thus, instead of the word 'sensorimotor' often used to describe epistemic activity that underwrites enactive perception, I suggest using the word 'motorsensory,' because it correctly captures the origin, causal direction, and sequence of epistemic actions: changes in one's proprioceptive states cause changes in one's exteroceptors and not vice versa.

Model of Enactive Visual Perception as Motorsensory Coordination Realized in a Special Kind of Neural Mapping

If the account of visual perception by O'Regan, Noe, Hurley and others is correct, it appears that acquisition of enactive visual perception necessarily involves a rather close correlation of instances of movement, initiated by the cognitive agent's motor system, with the resulting changes in one's visual field. Such a correlation, can from a neuroscientific point of view be structurally and functionally achieved by some kind of neural mapping between the regions of the brain that realize representations of proprioception (motor cortex and cerebellum) and the parts of the brain that realize visual representations (areas V1, V2, V3, MT, etc.).

In terms of content-vehicle analysis, one's engagement in epistemic actions produces correlated contents in one's proprioception and visual exteroception. Those correlated contents give rise to the correlation of their corresponding neural vehicles, and this correlation is stored in memory via Hebbian-type learning as the output-input mastery. The effect of such acquired correlation is that after repeated epistemic interaction, the agent's exposure to certain visual exteroceptive contents will activate related proprioceptive vehicles, thus making proprioceptive contents poised between the motor control structures and passive visual representations of the external reality for possible engagement in pragmatic actions (input-output mastery).

The outlined model gives us a structure for the analysis of complex epistemic and pragmatic activities in terms of the correlations of the multiple contents and mappings of the multiple vehicles of the modalities involved in any particular act of perception. I suggest that the application of the model of visual perception in terms of the proprioception-exteroception mapping should be extended to perception in other sensory modalities³. Thus, the thesis of enactive visual perception can be absorbed into a more general thesis of enactive perception stating that proprioception provides a unifying active platform onto which passive representations in all the sensory modalities are mapped. Furthermore, if perception in all the sensory modalities is proved to necessitate this kind of exteroception-proprioception mapping, one could insist that this particular neural organization is somehow implicated in the unity of human consciousness.

³ See Keeley (2001) for an account of what constitutes a sensory modality.

Specifying the Conditions for Falsification of the Thesis of Enactive Perception

The Tactile Vision Substitution and other similar systems⁴ can serve as a good testing platform for the thesis of embodied perception. However, before one can proceed with possible experimental designs, one needs to make sure that this thesis can be empirically tested in principle. That is, one must show that the thesis of enactive perception can be falsified. In other words, perception needs to be defined in such a way that it is logically independent of the motorsensory coordination and vice versa.

The model outlined in the previous section effectively restates the thesis of enactive perception in terms of motorsensory coordination and is going to be instrumental in the explication of the empirical nature of the thesis. Thus, the thesis of enactive perception stated in terms of a model of motorsensory coordination spelled out in propositional calculus looks like the following:

- (1) $\forall (x) (Px \supset Mx)$
 $Px = x$ has perception
 $Mx = x$ has motorsensory coordination

For any x , x has property P (perception) only if x has property M (motorsensory coordination). Therefore, for the enactive perception thesis to be empirically testable the conjunction of the following two statements has to be true⁵:

- (2a) $\exists (x) (Mx \cdot \neg Px)$

There exists such an x , which has property M (motorsensory coordination), but does not have property P (perception).

- (2b) $(x) (Px \cdot \neg Mx)$

⁴ See <http://www.seeingwithsound.com/sensub.htm> for a good overview of sensory substitution devices by Meijer.

⁵ Technically, 2b is sufficient to falsify the thesis, however 2a is important for my subsequent argument that proprioception is sufficient for perception.

There exists such an x , which has property P (perception), but does not have property M (motorsensory coordination).

In other words, we need to find at least one case in which a subject will have motorsensory coordination but no perception and, vice versa, a subject who will have perception but not motorsensory coordination. By identifying such cases one would make the motorsensory mapping logically independent of perception and perception logically independent of the motorsensory mapping.

Change blindness experiments are going to provide us with the instances corresponding to sentence 2a. Change blindness is a well-documented and studied form of perceptual invisibility: it could be very difficult to see a change in a scene, however, since once the change is attended to it becomes obvious. The change blindness effect can be directly experienced, for example, by visiting Ronald Rensink's University of British Columbia Visual Cognition Lab at <http://www.psych.ubc.ca/~viscoglab/demos.htm>. For our purposes one can modify the change blindness experiment and imagine that it involves an actual live scene instead of a picture and encourage a subject's motorsensory exploration of the scene after the change has taken place. The philosophical point of the change blindness effect is that in spite of the fact that the motorsensory mapping takes place as the subject visually examines the scene from different angles, she still does not have perception of the change.

A paramecium is a single cell organism that relies strictly on proprioception to navigate its environment. Its body's surface is lined with multiple hairs, which the paramecium uses for propulsion as well as identification and navigation around obstacles it encounters. Thus, when the paramecium bumps into an obstacle, the hairs reverse the direction of their vibrations, executing a maneuver to back up from the obstacle. Having reversed a sufficient distance, the paramecium executes a turn away and continues to swim forward. Thus, by relying on the proprioceptive stimuli arising within its own body, the paramecium successfully navigates its environment, that is, acquires perception of it.

To make the point more relevant to humans, one can imagine a person trying to navigate a maze with regular sensory organs of vision, audition, olfaction, gustation and touch temporarily disabled. With only proprioception at her disposal and the ability to report on her

progress (for our benefit) this person will be able to achieve perception of her environment by executing motor routines similar to those of paramecium: put one foot in front of the other until you no longer can, if you can't move your feet try turning right or left, or back out, then turn either way and start moving your feet forward again. By following this rather simple proprioceptive navigation routine a person will be able to successfully map out the maze and navigate it with ease. This thought experiment clearly demonstrates that perception is possible without the motorsensory mapping.

It appears that by executing sets of simple motor commands and, literally, probing its way around, any creature with proprioception can acquire an (admittedly rather crude) perception of its environment. In other words, proprioception on its own is necessary and sufficient for perception. Viewed from this perspective, visual perception can also be conceived of as a certain kind of "bumping" into things via encountering the reflections of light from environmental features. Thus construed, all sensory perception ultimately appears to be a multifaceted proprioceptive encounter with surroundings, albeit mediated by a variety of environmental phenomena like electromagnetic radiation, acoustic vibrations, distribution of chemicals in the air, etc.

Model of Perception as Enactive Motorsensory Coordination in Artificial Simulations

Having established the empirical nature of the thesis of enactive perception, it is illuminative to take a look at a couple of attempts to build perceptual robotic systems that implicitly or explicitly employ the motorsensory coordination model. The MIT's Humanoid Robotics Group at the MIT Artificial Intelligence Laboratory has developed several robotic platforms that provide a fertile ground for modeling and investigation of different features of the human mind, including experiments on enactive perception.

The Cog project from the outset was designed as an approximation of the motor and sensory dynamics of a human body. Cog's video cameras provide passive visual representations while joint positions and torque sensors enable proprioception, thus making it a very good platform for modeling perception as motorsensory coordination. In a series of experiments conducted by Metta and Fitzpatrick (2002), the basic ability

of Cog to perceive a cube in front of it was based on its initial epistemic interaction with the cube: probing and tapping to find its boundaries, manipulating the hand to measure the aperture of the cube in attempts to grasp it, etc. While Cog's hand was engaged in such epistemic interaction with the cube its cameras were following the changes of the visual field that Cog's movements provided. Simultaneously, the proprioceptive output was being mapped onto the visual input, thus enabling the robot to acquire perception autonomously by fusing vision and proprioception and later employing this in the pragmatic actions of grabbing and manipulating the cube.

Wermter et al. (2004) have been working on the Mirror-neuron Robot Agent (MIRA) with the goal of developing neural solutions for tasks that need to be solved by a robot that learns by the visually perceived demonstrations and verbal instructions. To enable perception in the MIRA the team of researchers used lateral associator connections between the neural networks that support areas of vision and proprioception of the robot. The weights of connections between representations in the visual net and the proprioception net develop concurrently, so that after repeated epistemic actions the robot acquires a background database of possible interactions with the target object. The epistemic actions of the MIRA robot resemble those of humans and other animals: approaching and backing away from the target object, lateral movements, poking and probing of the object and around it with a hand, etc. Eventually, repeated attempts and further fine-tuning of the visual-proprioceptive mapping result in the successful pragmatic action of grabbing the target object.

Setting the much debated issue of the qualia (Chalmers, 1995) aside, it is quite plausible that the robots in the experiments described above acquired perception of the target objects. Voluntary controlled action turns out to be an essential ingredient of unsupervised training of autonomous agents in the ecological context. While the cognitive agent's body provides a necessary reference frame for epistemic action and eventually grounds informative percepts to guide pragmatic action, it is the enactive motorsensory coordination that serves as the main mechanism that gets the job done.

Conclusion

All cognitive biological systems are embodied, and epistemic action is

an essential tool they have for recognizing and differentiating objects in the environment. Repeated engaging in motor actions (approaching, poking, pushing, etc.) toward an object reveals how its visual and other characteristics change when acted upon. Experiments with TVSS systems and robotic platforms demonstrate that, in a discovery mode, the visual (or pseudo-visual) system learns about the consequences of its motor actions and builds important motorsensory correlations. In a goal-directed mode the acquired motorsensory contingencies inform pragmatic actions by enabling the system to select the motor action that can cause a particular visual change. These two processes are inherently intertwined and together constitute a cognitive agent's ability to perceive objects in the environment.

The model of perception as the enactive motorsensory coordination gives us a useful structure for empirical analysis of complex epistemic and pragmatic activities in terms of the multiple contents and vehicles of the sensory modalities involved in any particular act of perception as modalities' correlated mappings onto proprioception. In particular, by providing the means to observe and reproduce the genesis of intentionality, i.e. awareness of something as external (the "appearance" of a phenomenon in a spatial perceptive field), the model makes it possible to conduct experimental studies in the area usually restricted to philosophical speculation. Furthermore, the model may provide a useful insight into the mechanisms underlying the apparent unity of consciousness by approaching the structure and functioning of consciousness in terms of the neural mappings of proprioception onto sensory modalities.

Further development of the thesis of enactive perception will involve the investigation into the nature of proprioception, because before a cognitive agent can engage in the active exploration of the environment she needs to map out the proprioceptive environment of her own body. In other words, a robot or a human needs to have her own proprioceptive map in place before proceeding to learn the motorsensory subject-object interactions. Another direction of interest would be investigation into whether the thesis of enactive perception can incorporate recent neuroscientific discovery by Fadiga et. al (2000) that implicates mirror neurons in the cognitive agent's understanding of the actions of others.

Finally, a larger area of the thesis' application would be the idea of the

embodied cognition put forth by Lakoff (1987), Johnson (1987), and Glenberg (1997). The rise of the notion of the embodied cognition reflects an increasing philosophical skepticism concerning the ultimate merit of the intuitive divisions of the mental processes into perception, action, and cognition. In other words, it appears that cognitive development cannot be usefully treated in isolation from issues concerning the cognitive agent's physical embedding in, and interaction with the environment. The embodied cognition involves perception, action, and thought as bound together in a variety of complex and interpenetrating ways. The outlined above enactive motorsensory coordination model can provide a foundation and a useful tool for developing structural and functional models of the embodied cognition.

Works Cited

- Bach y Rita, P. (1972). Brain mechanisms in sensory substitution. New York : Academic Press.
- Block, N. 2001. Behaviorism revisited. Behavioral & Brain Sciences 24: 977-978.
- Brooks, R.A., Intelligence without representation, Artificial Intelligence 47 (1991), 139-159.
- Chalmers, D., (1995). The conscious mind: in search of a fundamental theory, Oxford University Press, Oxford.
- Clark, A. (1997). Being There: Putting Brain, Body, and World Together Again. MIT Press.
- Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (2000). Visuomotor neurons: ambiguity of the discharge of 'motor' perception? International Journal of Psychophysiology, 35:165-177.
- Fitzpatrick, P. and Metta, G. "Towards Manipulation-Driven Vision", 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Lausanne, Switzerland, 2002.
- Gibson, J.J., (1979). The Ecological Approach to Visual Perception. Houghton Mifflin.

- Glenberg, A.M. (1997). What memory is for. Behavioral and Brain Sciences 20 (1): 1-55.
- Johnson, M. (1987). The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. Chicago: The University of Chicago Press.
- Hurley, S. L. (1998). Consciousness in Action. Cambridge, MA: Harvard University Press.
- Keeley B. L. (2001). Making Sense of the Senses: Individuating Modalities in Humans and Other Animals, Journal of Philosophy, April 2001.
- Kirsh, D. & Maglio, P. (1994). On distinction between epistemic from pragmatic action. Cognitive Science, 18, 513-549.
- Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.
- Marr, D. (1982). Vision: a computational investigation into the human representation and processing of visual information. New York: W.H. Freeman and Sons.
- Merleau-Ponty, M. (1945). The Phenomenology of Perception. Translated by Colin Smith, London: Routledge 2002.
- Metta, G. and Fitzpatrick, P. "Better Vision through Experimental Manipulation", In EPSRC/BBSRC International Workshop, Biologically-Inspired Robotics: The Legacy of W. Grey Walter, Bristol, UK, August 2002.
- Noe, A. (2004). Action in Perception. MIT Press.
- O'Regan, J. K. and A. Noe. (2001) A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24, no 5: 939-973 Cambridge University Press.
- Rensink, R.A., J.K. O'Regan, and J.J. Clark. (2000). On the failure to detect changes in scenes across brief interruptions. Visual Cognition 7, no. 1: 127-146.

Sacks, O. (1995). An Anthropologist on Mars: Seven Paradoxical Tales. New York: Knopf.

Skinner, B.F. (1957). Verbal Learning. New York: Appleton-Century-Crofts.

Varela, F., Thompson E., and Rosch. E. (1991). The Embodied Mind: Cognitive Science and Human Experience. MIT Press.

Wermter S., Weber C., Elshaw M., Panchev C., Erwin H., Pulvermüller F. (2004). Towards multimodal neural robot learning in Robotics and Autonomous Systems 47, 171-175.

FROM LOGICAL TO EPISTEMIC CIRCULARITY: THE CARTESIAN CIRCLE REMAINS

TAN PAI-EN

National University of Singapore

Introduction

One of the most important and perhaps the most difficult question in epistemology is the problem of the criterion. The Cartesian Circle can be seen as a particular case of this problem. In this paper I look at a well-known attempt by James van Cleve to solve the Cartesian Circle by distinguishing between knowledge falling under some criterion and knowledge of that criterion. I will argue that van Cleve's procedure fails because it is epistemically circular and highlight the problem with epistemically circular arguments. I will conclude that, since epistemically circular arguments do not sufficiently answer skeptical considerations and hence do not provide a solution to the problem of the criterion, the Cartesian Circle remains.

The problem of the criterion is characterized by Roderick Chisholm as the difficulty in answering the following two questions in epistemology:¹

α. What do we know?

β. What are the criteria of knowledge?

It appears, *prima facie*, that we cannot answer *α* unless we know the answer to *β*, and conversely, that we cannot answer *β* unless we know the answer to *α*.

In 'Foundationalism, Epistemic Principles and the Cartesian Circle,' James van Cleve responds to the problem of the criterion with a solution to the Cartesian Circle which makes use of the distinction between different levels of knowledge: knowledge that falls under a criterion and knowledge of the criterion. Van Cleve suggests that this solution to the Cartesian Circle can be used to solve the problem of the criterion as well. I shall first examine van Cleve's solution in detail, and highlight the problems with his solution thereafter.

¹ Chisholm, *The Foundations of Knowing*, 65.

Van Cleve's Solution to the Cartesian Circle

Van Cleve spells out the apparent problem of the Cartesian Circle as follows:²

(I) I can know (be certain) that (p) whatever I perceive clearly and distinctly is true only if I first know (am certain) that (q) God exists and is not a deceiver.

(II) I can know (be certain) that (q) God exists and is not a deceiver only if I first know (am certain) that (p) whatever I perceive clearly and distinctly is true.

If the terms in (I) and (II) mean the same thing, and both (I) and (II) are true, then, as van Cleve says, "I can never be certain of either p or q . Yet Descartes seemed to certain of both p and q . Thus, any solution to the Cartesian Circle will have to deny either (I) or (II), or make a distinction in the meaning of their words. Van Cleve's solution is to deny (II).

To appreciate his solution, we need to recognize a subtle distinction that van Cleve points out. Consider the following:

- (A) For all P, if I clearly and distinctly perceive that P, then I am certain that P.*
(B) I am certain that (for all P, if I clearly and distinctly perceive that P, then P).

As van Cleve explains, (A) says that whenever I clearly and distinctly perceive any perception, then I am certain of it, whereas (B) is a "general principle connecting clear and distinct perceptions with truth"³ Clearly, (A) could be true even if (B) were not. Such a case would be when (A) is true and a subject is unaware of (A)'s truth.

Van Cleve proposes that (A) serve as an epistemic principle. Now, given that (A) is true, Descartes does not need to know (A) before he proves God's existence and His non-deceiving nature in order to have certainty. Descartes need only fall under (A). This is the insight that van Cleve's solution rests upon. According to (A), the antecedent (clear and distinct perceptions) is sufficient for certainty, hence nothing else is necessary. Descartes does not even need to have the concept of clear and distinct perceptions, or the awareness that he is having such perceptions. He need only have clear and distinct perceptions per se. Thus, Descartes starts off with only (A) being true (but not known), and arrives at (B)

² van Cleve, "Foundationalism, Epistemic Principles and the Cartesian Circle," 55.

³ *Ibid.*, 67.

only after (A) is proved. Therefore (II) (the lower arc of the Circle) is denied, because to be certain that God exists and is not a deceiver does not require that “I first know that (whatever I perceive clearly and distinctly is true),” which is (B), but only that “Whatever I perceive clearly and distinctly is true,” which is (A).

Since we are going to examine van Cleve’s solution in detail, it is helpful to reproduce the stages of the Cartesian enterprise as he conceives of it:⁴

Stages	Propositions	Description
1	<i>I think, the causal maxims, etc.</i>	Propositions known because they are clearly and distinctly perceived.
2	<i>God exists, God is no deceiver.</i>	Propositions known because they are clearly and distinctly perceived to follow from premises at stage 1.
3	<i>Whatever I perceive clearly and distinctly is true.</i>	Principle known because it is clearly and distinctly perceived to follow from propositions at stage 2.
4	<i>I perceive clearly and distinctly that I think, etc.</i>	New premises, one corresponding to each premise at stage 1.
5	<i>I am certain that I think, etc.</i>	Propositions known because they are clearly and distinctly perceived to follow from propositions at stages 3 and 4.

According to van Cleve, Descartes becomes certain of the propositions above in the order from stage 1 to stage 5. At stage 1, Descartes has no supporting grounds (known to himself) for the propositions, but they serve as his premises for proving (A). But this does not matter - as long as (A) is true, the clear and distinct propositions at stages 1 and 2 are “immediately justified,”⁵ and Descartes can move on to stage 2 and 3. Once stage 3 is reached, (A) is proved and provides an answer to the objection that the initial premises are arbitrary. In other words, Descartes

⁴ Ibid., 72.

⁵ Ibid., 73.

“initial premises are immediately justified and certain, but the higher-order proposition that says that they are certain is justified by appeal to reasons.”⁶ The higher-order proposition, of course is (B).

Once (A) is proved, Descartes can then move on to propositions in stages 4 and 5. He can then apply (A) to any clear and distinct perception and obtain justification for it.

Although van Cleve’s solution may seem somewhat circular, we cannot charge him with making a logically circular argument, because at no stage is (A) actually used in the argument. If there is any circularity, it is not logical in nature.

Objections to van Cleve’s Solution to the Cartesian Circle

Having outlined van Cleve’s argument, I will now present two objections. The first objection concerns the obtaining of metaphysical certainty over psychological certainty, and the second objection aims to show that there is a fundamental problem with arguing as van Cleve does – namely, arguing to justify epistemically circular arguments. To be exact, there is a problem with arguing for an epistemically circular argument as the answer to the problem of the criterion. As it turns out, these two objections are not independent – the first is in fact just a particular way of pointing out the general problem highlighted by the second.

First Objection: Metaphysical Certainty versus Psychological Certainty

It appears that van Cleve does not take into account the implications of the psychological nature of the clear and distinct perceptions that he considers at stages 1 (and 2) on the status of the (A) which is arrived at stage 3 as a result. What exactly is the status of those propositions at stage 1? Van Cleve characterizes them as “propositions known because they are clearly and distinctly perceived.”⁷ We must be careful here what he means by this. At the outset, we can say at least that for a proposition P at stage 1, a subject S cannot assert “I clearly and distinctly perceive P”, because such propositions belong to stage 4. So even though S clearly and distinctly perceives that “I think,” he cannot assert this fact himself. Not only does S not know (A), he does not even know “I clearly and

⁶ Ibid., 73.

⁷ Ibid., 72.

distinctly perceive that I think, etc.”⁸ But since van Cleve claims that S is nevertheless certain of propositions at stage 1, I take this to mean that stage 1 is characterized by some sort of immediate knowledge that is free of any sort of reflection, recollection or higher-order certainty. In other words, S is justified at stage 1 simply because of immediate clear and distinct perceptions, because that is all (A) requires for its antecedent.

A problem arises because the kind of certainty that S has of propositions at stage 1 seem to be similar to the kind of psychological certainty that Alan Gewirth⁹ begins with in his argument which van Cleve criticizes. If S cannot even know “I clearly and distinctly perceive that I think”, but merely has an immediate perception which is clear and distinct, it seems that the certainty is only psychological. In other words, the subject is certain only insofar as there is a psychological compulsion to believe due to clear and distinct perceptions. Certainly S has no epistemic justification which he can apply to *himself* at this point. Furthermore, the view that Descartes is concerned with psychological certainty has been plausibly argued for by others such as Louis E. Loeb¹⁰ and has much textual evidence. For example, in paragraph 4 of the Third Meditation, Descartes writes,

Yet when I turn to the things themselves which I think I perceive very clearly, I am so convinced by them that I spontaneously declare: let whoever can do so deceive me, he will never bring it about that I am nothing, so long as I continue to think I am something:..... or bring it about that two and three added together are more or less than five, or anything of this kind in which I see a manifest contradiction.¹¹

It is only *after* this passage that Descartes says that any reason to doubt the above is “so to speak, a metaphysical one,” and that in order to remove any such doubt,

I must examine whether there is a God, and, if there is, whether he can be a deceiver. *For if I do not know this, it seems that I can never be quite certain of anything else.* [italics mine]¹²

⁸ *Ibid.*, 72.

⁹ Gewirth’s arguments are taken from the papers by him on the Cartesian Circle, namely: “The Cartesian Circle,” “The Cartesian Circle Reconsidered,” and “Descartes: Two Disputed Questions,”

¹⁰ Louis E. Loeb, “The Cartesian Circle,” 200-235. Loeb actually argues for a full psychological interpretation, meaning that even at the end Descartes achieves only psychological certainty and not epistemic or metaphysical certainty, but the point still holds for propositions held before God’s existence and non-deceiving nature is proved.

¹¹ Descartes, *Meditations on First Philosophy*, 25.

¹² *Ibid.*, 25.

Clearly, if the kind of certainty achieved in the beginning is metaphysical certainty, then Descartes would not bring in the skeptical doubt after talking about such certainty. And in paragraph 14 of the Fifth Meditation, he writes

Admittedly my nature is such that so long as I perceive something very clearly and distinctly I cannot but believe it to be true.... For example, when I consider the nature of a triangle, it appears most evident to me... that its three angles are equal to two right angles; and so long as I attend to the proof, I cannot but believe this to be true.¹³

Indeed, the above passages suggest that there is something irresistible about the nature of immediate clear and distinct perceptions such that it is impossible not to be convinced of them. (By *immediate*, I mean not only that the clear and distinct perception is current, but also that it is of the type of proposition belonging to stage 1 (eg. "I exist"), and *not* stage (4) (eg. "I clearly and distinctly perceive that I exist."). The irresistibility applies to the former but not necessarily to the latter. Descartes' point is that once we turn to other propositions, such as that of a deceiving God, then our certainty is shaken. This happens because according to him we cannot fix our mental vision continually on the same thing,¹⁴ and when once we stop attending to those clear and distinct perceptions, a proposition such as that of the deceiving God (or evil demon, etc.) constitutes a skeptical hypothesis that serves as a defeater for our certainty in those clear and distinct perceptions. As such, Descartes seeks to remove the possibility of such skeptical hypotheses by proving God's existence and His non-deceiving nature. But since he starts off with psychological certainty, the only resources available for the removal of metaphysical doubt are psychological. On Descartes' view, therefore, psychologically certain propositions must be sufficient for the removal of metaphysical doubt.

Earlier in his paper, van Cleve criticizes Gewirth's definition of metaphysical certainty. On Gewirth's own reading of Descartes, he concludes that "there can be clearness and distinctness without truth."¹⁵ Clearly, he means that psychological certainty about the obtainment of a state of affairs can be present even if that state of affairs did not obtain. So far, so good. But the important question is how we move on to the next step. Here, the notion of metaphysical certainty is introduced. According to van Cleve, Gewirth holds that "a proposition P is metaphysically certain if and only if there is no proposition R that

¹³ *Ibid.*, 48.

¹⁴ *Ibid.*, 48.

¹⁵ Gewirth, "The Cartesian Circle," 373. 58.

is a reason for doubting P,"¹⁶ and that P starts off only psychologically certain. The metaphysical certainty is achieved when every skeptical hypothesis against P (e.g. the deceiving god) is removed. Gewirth justifies this claim by explaining that "any validity which [a] hypothesis may possess can be only of the sort permitted by the method in accordance with which it has been erected."¹⁷ In other words, since the doubt itself is supposedly raised in the realm of psychological certainty, for which 'clearness and distinctness' is the norm for certainty, all we need is a clear and distinct argument that removes the skeptical hypothesis. Now it is not clear that the preceding statement is true, since it is doubtful that the metaphysical doubt is raised in the realm of psychological certainty. Rather, it is precisely the kind of psychologically irresistibility to assent to whatever is clear and distinct while attending to it that the metaphysical doubt challenges. The irresistibility applies only to the clear and distinct perceptions, and not to the metaphysical doubt. But even if we ignore the abovementioned objection, Gewirth still falls prey to van Cleve's claim that he has "merely extended the class of psychological certainties"¹⁸ by virtue of the way that metaphysical certainty is defined by the latter, meaning that we are now also psychologically certain of the falsehood of every reason for doubting P (and not just P itself). I think that van Cleve is right here. Although the validity of the skeptical hypothesis is psychologically destroyed, that says nothing about its metaphysical status. Indeed, following Gewirth's own claim about the validity of any hypothesis being limited by the method with which it has been erected, we may plausibly infer only the psychological certainty of the removal of metaphysical doubt. If that is what metaphysical certainty amounts to, then any gain we have made over merely psychological certainty *per se* seems trivial (i.e. we now have psychological certainty of our psychological certainty).

Now compare this with van Cleve's solution. Presumably, van Cleve will point out that the kind of certainty that S has at stage 1 is more than psychological certainty, because by virtue of (A), metaphysical certainty is already conferred, even though S does not know (A). As long as propositions are clear and distinct, they fall under (A) and that is sufficient for van Cleve. While this is true if (A) is true, it remains the case that S can only work with psychological certainty before (A) is proved for S himself. If S possesses any metaphysical

¹⁶ van Cleve, "Foundationalism, Epistemic Principles, and the Cartesian Circle," 61; van Cleve, "Foundationalism, Epistemic Principles, and the Cartesian Circle," 58.

¹⁷ Gewirth, "The Cartesian Circle," 391.

¹⁸ van Cleve, "Foundationalism, Epistemic Principles, and the Cartesian Circle," 61; Gewirth, "The Cartesian Circle," 391.

certainty before stage (3), S is not aware of it. But that means that at stage 1, psychologically clear and distinct perceptions are sufficient for metaphysical certainty. In other words, the subjective, psychological certainty of S is rendered objectively and metaphysically certain by virtue of (A). Van Cleve makes no mention of the psychological aspect, but I think that the compelling textual evidence, together with his characterization of propositions at stage 1, warrants such an account. If so, is there a problem for van Cleve?

What (A) does at stage 1 is to confer metaphysical certainty upon immediate clear and distinct perceptions which we have shown to be merely psychologically certain from the vantage point of the subject. But how is this principle (A), laden with such powers of metaphysical bestowal, arrived at? By virtue of the familiar route whereby Descartes rules out the idea of a deceiving God, and so on – which is precisely the elimination of any doubt regarding clear and distinct perceptions themselves. In other words, (A) is the conclusion of an argument which proceeds by eliminating doubt of clear and distinct perceptions, by relying on clear and distinct perceptions themselves. In that case, is (A) psychologically or metaphysically certain? It is hard to see how different this is from Gewirth's argument, insofar as we are taking the point of view of the subject S. Van Cleve will no doubt point out the obvious difference that Gewirth's type of argument begins with psychological certainty, while his procedure starts off already with metaphysical certainty. Taken this way, S has metaphysical certainty at every stage. But van Cleve's task is to justify the procedure of the subject in question, and the metaphysical doubt arises in S precisely because he cannot claim metaphysical certainty for himself to begin with. Thus proceeding, S cannot claim metaphysical certainty of (A) either. Of course, (A) *could* be metaphysically certain, but *is* it? And how would the subject ever come to know that?

Even so, in defense of van Cleve, he notes elsewhere that the rule of clearness and distinctness is a rule involving evidence¹⁹, and hence we may suggest that there is more going for (A) than Gewirth's definition of metaphysical certainty. Descartes himself refers to clear and distinct perceptions as "matters in which I believe myself to have the best evidence."²⁰ Since the rule of clearness and distinctness is equivalent to principle (A), does principle (A) contain more than just the removal of the possibility of doubt regarding psychological certainty? I think it does, and that van Cleve is right here, because Gewirth's argument

¹⁹ *Ibid.*, 71.

²⁰ Descartes, *The Philosophical Writings of Descartes Vol. II*, 24.

for metaphysical certainty seems to rest solely on the removal of any proposition R for doubting P. But principle (A) is arrived at through the removal of any reason for doubting “clear and distinct perceptions”, which may be certain for other reasons, such as external or separate evidence.²¹ The fact that immediate clear and distinct perceptions are psychological in nature does not preclude the possibility that these perception are reliable; one may, for instance, give arguments for the proper functioning of cognitive faculties. Externalist forms of justifications such as these may not be available to Descartes if he is an internalist (as commonly thought), but van Cleve’s solution actually has an externalist bent, because at stages 1 and 2, the justification for S does not come from any sort of epistemic or doxastic duty performed through reflection on available evidence or right judgment with any knowledge of an epistemic principle like (A), but rather by virtue of (A) being true. And following van Cleve’s distinction, the truth of (A) is an external, objective notion without anything whatsoever to do with S’s knowledge of it. This, together with the irresistibility of the psychological nature of those basic propositions, makes it hard for S to perform any conscious epistemic duty at the initial stages. (A) becomes a kind of meta-justification for the initial premises, and thus we have an externalist foundationalism at this point.

Having said this, van Cleve must then face the problem of finding some external form of justification for clear and distinct perceptions such that the removal of them would suffice for metaphysical certainty. But even if this were possible, it appears that we would be moving further from Descartes’ original enterprise rather than closer to it. It would suffice to note here that it is far from easy to find a convincing externalist form of justification that is not epistemically circular, and in the next objection I shall point out what is wrong with epistemic circularity.

Second Objection: Circularity and Justification in Van Cleve’s Solution

This brings us to my second objection. The crucial question to ask of van Cleve’s solution is whether it is circular or unjustified. I will argue that van Cleve cannot escape both charges at the same time.

Now suppose (A) is true. Van Cleve starts with a set of propositions known to the subject S (stage 1 and 2) because they fall under (A). At this stage S does not know, and has no grounds for believing, that basket of 21 Gewirth, of course, starts off with clear and distinct perceptions in his own argument also. But his definition of metaphysical certainty, according to van Cleve, is such that the ‘absence of any proposition R for doubting P’ is sufficient for such certainty. Principle (A), on the other hand, requires specifically “clear and distinct perceptions.”

propositions to be true or certain. If S were asked why he holds these propositions, S could not reply with principle (A) since (A) is not known at this time.

Now, if S were asked at this stage to justify his premises, what could he reply with? Van Cleve will no doubt reply that after (A) is proved, that is not a problem. According to him, Descartes can reply as follows: "Those premises are things I knew for certain. The proof of this is that I perceived them clearly and distinctly, and whatever I so perceive is certain."²² Is that really a satisfactory answer to our question? In order for an argument to be sound, at least the premises must be true, or generally accepted as certain. When the truth of the premises in 1 and 2 is questioned, the only possible reply is from the conclusion (A). Thus, while (A) is not used in the argument itself, there is a wider circularity that results when one defends the argument against the charge that its premises are unjustified. Clearly, the conclusion is necessary for justification of premises. Thus, the argument itself may not be logically circular, since (A) is not used or implicit in the premises, but surely there is epistemic circularity in the sense that the argument "involves a commitment to the conclusion as a presupposition of our taking ourselves to be justified in accepting the premises"²³

On the face of it, it may appear that epistemic circularity is not vicious, because if (A) is true, S can have certainty regardless of whether (B) is true. As van Cleve notes, for an epistemic principle "If [...] then p is justified for S', the obtaining of whatever condition specified in the antecedent is sufficient for p's being justified for S."²⁴ And "if X is sufficient for Y, then there is no other condition Z that is necessary for Y unless Z is also necessary for X."²⁵ Thus van Cleve wants to make the general point that epistemic principles do not need to be known in order for knowledge to arise from them. So, if S has certainty from (A), assuming he is able to prove (A), he attains (B), which is the knowledge of (A) as a general epistemic principle. Such epistemic circularity in achieving justification may sound acceptable, especially considering that the alternatives seem to be arbitrariness or an infinite regress of justifying propositions.

We can certainly think of examples where a person S does not know (A) but satisfies (A). But that is because *we* (and not S) already suppose

22 van Cleve, "Foundationalism, Epistemic Principles, and the Cartesian Circle," 72.

23 *Ibid.*, 72.

24 *Ibid.*, 77.

25 *Ibid.*, 77.

the truth of (A) - that is, in our example we are implicitly supposing ourselves to know, on behalf of S, that (A) is true. And that is none other than (B). In other words, to accept van Cleve's procedure, we must have knowledge that the subject S does not possess. While S does not know (B), we do. But the fact of the matter is, in the actual world, following van Cleve's procedure, none of us has access to that second-order, reflective knowledge at all. All we have are propositions that are certain if some epistemic principle (A) is true. But we never know whether there exists a principle like (A) which is true, and hence we cannot be certain of those propositions. While those propositions may be clear and distinct, we cannot have reflective knowledge of the certainty of our propositions, because we have potential defeaters of the certainty such as the evil demon, living in the Matrix, and so on. And without this reflective knowledge - knowing that we are certain of some proposition - we cannot move on to deduce other propositions (such as "God exists and is not a deceiver") with reflective certainty. And hence we cannot reach (B), which is reflective, second-order knowledge of (A) itself. This means that we are in the same position as S is in.

All the while, in considering van Cleve's argument, we have put ourselves on the level of objective knowledge - i.e. knowing that (A) is really true - while S has only subjectively clear and distinct perceptions. It is we who hypothetically accept that S is justified at stages 1 and 2, who confer that metaphysical certainty upon S's clear and distinct perceptions. But in actuality none of us has any epistemic advantage over the imaginary person S, or any real person. Insofar as Descartes' epistemological project is concerned, none of us actually starts from the position of knowing (A) or any epistemic principle to be true. (If we did know (A), then we would be immune to the evil demon.) And since we do not actually know (A) to be true, our accepting that S is justified in believing those clear and distinct perceptions at stages 1 and 2 has no justification. Therefore, the circularity arises in assuming that objective epistemic position for ourselves when in reality we ourselves start in the same subjective position as S. As we do not actually have this sort of access to some objective epistemic principle (following van Cleve's sort of argument), the circularity invalidates the argument.

The problem may be clearer if we present the situation like this: van Cleve wants to argue for a solution of the Cartesian circle. In other words, his paper is an argument for the soundness of Descartes' argument (or procedure) in the Third Meditation. Descartes' argument is presented as follows: Let P stand for the propositions at stage 1 (I think, I exist, causal maxims) and Q for the propositions at stage 2 (God exists,

is not a deceiver). Then we have the following (ignoring the problems with Descartes proof of God and His non-deceiving nature and their implications):

- D1. *P.*
- D2. *If P, then Q.*
- D3. *If Q, then (A).*
- D4. *Therefore, (A). (D2, D3)*

The argument above is obviously valid. But in order for it to be sound, we need to examine the premises themselves. In the present case, it seems clear to me that we *have* to justify P, because of the doubts raised in the First Meditation, particularly with regards to the evil demon. Van Cleve seems to agree, because he claims that P is justified by (A). Now, how exactly are we to formulate van Cleve's argument? What does he need in order to convince *us* that P? First, he would need to affirm that "If (A), then P", and secondly, he would need to claim that (A) is true (although S does not know this initially). However, there is no argument for (A) that van Cleve provides other than the one Descartes gives. As such, it is difficult to see what van Cleve's argument amounts to other than this:

- V1. *If (A), then P.*
- V2. *(A).*
- V3. *Therefore, P.*

Argument for V2:

- W1. *If P, then (A).*
- W2. *P.*
- W3. *Therefore, (A).*

Now it is clear what the problem is. We need to justify P. To do that, we need (A). But to accept (A), we need P itself. Thus, while Descartes' argument *per se* as presented by van Cleve is unjustified, it can be justified by van Cleve's larger argument. But to do so, van Cleve's larger argument must be circular.

Someone may still object by saying: "If we don't stop somewhere in our quest for justification, then we will need an infinite number of justifying propositions. (A) seems intuitive enough for me, so why not start with it? Why should the burden of proof lie on van Cleve to justify (A)?" Whether it is reasonable or not to demand justification for premises that seem intuitive and where the burden

of proof actually lies in general cases of philosophical arguments are meta-philosophical questions that are not simple. However, it is clear that the epistemological problem is phrased in the context of the *Meditations*, against the backdrop of skepticism.

All in all, what we really want to know is more than an epistemic principle like (A). In philosophy, and particularly in epistemology, we want to know "Why (A)?" We want some justification for (A) that does not rest on (A) itself. If I inquire why (A) is true, I do not expect (A) to be used to resolve my doubt. But the call for justification can only be answered on pains of circularity. As Reid puts it,

If a man's honesty were called in question, it would be ridiculous to refer it to the man's own word, whether he be honest or not. The same absurdity there is in attempting to prove, by any kind of reasoning [...] that our reason is not fallacious, since the very point in question is, whether reasoning may be trusted.²⁶

Can van Cleve say anything in reply? He may repeat that we do not need to know (A) at all, because (A) can still be objectively true. Clearly then, (A) remains unjustified. One can easily assert "not-(A)" at the beginning, and there will be no answer to give, and no argument with which to proceed.

Some may want to insist that there is nothing wrong about this sort of circularity. After all, part of what van Cleve (or Descartes) seeks – and part of the quest of epistemology – is to know how one can know anything. And admittedly, van Cleve does offer reasons for believing that one can know something with his theory. But the problem is that there are no reasons offered for the theory itself. In other words, if (A) is true, then van Cleve is correct. But we have no reason for accepting (A) apart from depending on (A) itself. Any justification of van Cleve's procedure is thereby logically circular, and thus we have no good reasons to accept the procedure itself. Consider Stroud's assessment of epistemic circularity:

[The externalist epistemologist] [...] is at best in the position of someone who has good reason to believe his theory if that theory is in fact true, but has no such reason to believe it if some other theory is true instead. He can see what he would have good reason to believe if the theory he believes were true, but he cannot see or understand himself as knowing or having

²⁶ Reid, *Inquiry and Essays*, 276.

good reason to believe what his theory says.²⁷

What Alston says about epistemic circularity (with respect to the reliability of sense perception) applies here as well:

Epistemic circularity does not in and of itself disqualify the argument. But even granting that point, the argument will not do its job unless we *are* justified in accepting its premises; and that is the case only if sense perception is in fact reliable [or if (A) were true, for our case].²⁸

Alston goes on to point out that such circularity makes us helpless in discriminating between those belief-forming practices which are reliable and those which are not.

We can say the same of any belief forming practice whatever, no matter how disreputable. We can just as well say of crystal ball gazing that if it *is* reliable, we can use a track record argument to show that it is reliable. But when we ask whether one or another source of belief is reliable, we are interested in *discriminating* those that can reasonably be trusted from those that cannot. Hence merely showing that *if* a given source is reliable it can be shown by its record to be reliable, does nothing to indicate that the source belongs with the sheep rather than with the goats [...]. Hence I shall disqualify epistemically circular arguments on the grounds that they do not serve to discriminate between reliable and unreliable doxastic practices.²⁹

What Alston says here applies equally for principle (A) as for sense perceptions, because the derivation of principle (A) relies on its own support. However, there is a possible rejoinder to such objections.³⁰ One may reply that the choosing of any arbitrary principle (K) does not guarantee that there is any argument whose premises are justified by (K) *and* which leads to conclusion (K) itself. In other words, van Cleve's procedure does not depend solely on principle (A) supporting itself; it also makes use of an argument whose premises are justified by (A) to achieve the conclusion of (A) itself.

But if we grant this, then we have the difficult job of showing exactly how Descartes' arguments for principle (A) actually work. Descartes' arguments for God's existence and His non-deceiving nature are dubious, and hence we have no guarantee that principle (A) *is* actually

27 Stroud, "Understanding Human Knowledge in General," 43.

28 Alston, *The Reliability of Sense Perception*, 17.

29 *Ibid.*, 17.

30 For example, see Frederick F. Schmitt, "What's Wrong with Epistemic Circularity?," 385.

the conclusion of a valid argument. Since van Cleve is only concerned with whether Descartes' general procedure is circular, he does not attempt to deal with the problems with the actual arguments that lead to principle (A). Hence we do not know that Descartes' argument for (A) is valid, and the rejoinder does not work here. But, more importantly, as long as it is *possible* for some principle (K) which represents dubious belief-forming practices to work like (A), it is sufficient for the preceding objection to stand.

Finally, consider the context of the *Meditations*. Descartes himself stated that the knowledge of a philosopher should be "so firm that it is impossible for us ever to have any reason for doubting,"³¹ so strong that it can never be shaken."³² As Newman and Nelson have noted,³³ Descartes' work came after a revival of Greek skepticism, where the writings of Pyrrhonian and Academic skeptics, as well as contemporary skeptics like Montaigne, were prominent. Descartes claims that he had "long ago seen several books written by the Academics and Skeptics,"³⁴ and in the Seventh Replies, Descartes refers to "the arguments by means of which I became the first philosopher ever to overturn the doubt of the skeptics."³⁵ Therefore, any attempt to secure the kind of knowledge that Descartes aims for faces the challenge of answering the Pyrrhonian doubt, as outlined by Sextus Empiricus himself:

[...] to decide the dispute that has arisen about the criterion, we have need to an agreed-upon criterion by means of which we shall decide it; and in order to have an agreed-upon criterion, it is necessary first to have decided the dispute about the criterion. Thus, with the reasoning falling into the circularity mode, finding a criterion becomes aporetic; *for we do not allow them to adopt a criterion hypothetically*, and if they wish to decide about the criterion by means of a criterion we force them into an infinite regress [italics mine].³⁶

It is not hard to see that the underlined part is none other than the problem of the criterion. Clearly then, the skeptic of Descartes time would demand justification for a principle like (A), and the hypothetical adoption of a criterion like (A) is explicitly forbidden by Empiricus. Lammaranta's remarks on the reliabilist apply equally to Descartes, as

31 Descartes, *The Philosophical Writings of Descartes Vol. II*, 103.

32 Descartes, *The Correspondance*, 147.

33 Newman and Nelson, "Circumventing Cartesian Circles," 370.

34 Descartes, *Oeuvres De Descartes Vol. VII*, 130.

35 Descartes, *The Philosophical Writings of Descartes Vol. II*, 376.

36 Sextus Empiricus, *The Skeptic Way: Sextus Empiricus's Outlines of Pyrrhonism*, II 4.

both would think that

[...] it is an open question whether the process [of reasoning according to clear and distinct perceptions] is reliable or not. For [them] it is thus an open question whether the premises of the argument are justified and so also whether it justifies the conclusion.³⁷

The hypothetical holding of the criterion (A) before (A) is proved would be unacceptable. Thus, it is unlikely that Descartes himself, who was surely aware of this, would have endorsed van Cleve's interpretation. After all, we are not only interested in knowing how we *could* have knowledge (which an epistemically circular account of knowledge does arguably answer), but whether we *do* have knowledge at all.

In summary, the epistemic circularity exhibited by van Cleve's procedure is unacceptable because of the following:

1. *It is helpless against a claim that principle (A) is false;*
2. *Either Descartes' premises or (A) needs to be justified, but neither can be done without giving a circular argument; and*
3. *It makes such a procedure indiscriminate from unreliable methods of knowledge acquisition which are epistemically circular.*

These points are sufficient to provide the skeptics of Descartes' time with enough ammunition to destroy the project for knowledge that is indefeasible. Any philosophical argument that suffers from such defects would be unacceptable, and I suspect that the reason why theories in epistemology which exhibit such circularity are still accepted by some philosophers is simply the lack of better non-circular forms of justification. Alternatively, one might suggest that any search for justification in the spirit of Descartes' enterprise is misguided in the first place.

Conclusion

I have presented two objections. In the first objection I have attempted to show that both van Cleve's description and textual considerations demand a psychological interpretation at least for the propositions at stage 1. This brings van Cleve under his criticisms of Gewirth, the escape from which leads us to require an externalist justification for clear and

³⁷ Lammenranta, "Reliabilism and Circularity", 115.

distinct perceptions. Whether an externalist interpretation of Descartes is acceptable (at least for stages 1 and 2) is doubtful. There is no room for further discussion, but at the very least it seems that Descartes' notion of justification has something deontological about it,³⁸ while the justification at stage 1 is not really up to the subject's will. Furthermore, it is hard to find externalist justification that does not suffer from similar forms of epistemic circularity as the kind pointed out in the second objection.³⁹

The second objection highlights the fact that we have no epistemic advantage over S, and therefore we are not in the position to accept van Cleve's argument without having to either 1) leave the premises of the argument or (A) unjustified against the skeptical hypotheses or 2) incur logical circularity in defending the premises of the argument. This is also why van Cleve cannot simply claim that S has metaphysical certainty by virtue of (A) – which is what the first objection drives at. As a result, van Cleve's argument does not sufficiently answer the skeptical considerations.

Although this paper is concerned largely with van Cleve's solution of the Cartesian circle, its point of interest also lies in the fact that the objections to epistemic circularity also apply to other common considered sources of knowledge. The reliability of memory, sense perception, deduction, induction, and so on, stands in need of justification if we are to answer the skeptics convincingly. Yet widely-accepted theories of justification for these sources which are not epistemically circular have not emerged. Although some have argued that such circularity is acceptable, such arguments tend to be claims that fulfillment of the epistemic principle is sufficient, and justification of theory or procedure itself is unnecessary.⁴⁰ To this, I can only reply that such sufficiency is not what the skeptics were primarily out to undermine in the first place, and it is cold comfort that we can settle for this. By normal standards, any philosophically satisfactory account of knowledge should *itself* be convincing. If my objections bear weight, then it seems that we do not have convincing arguments against skepticism.

38 In Med IV, paragraph 12, Descartes remarks: "If, however, I simply refrain from making a judgement in cases where I do not perceive the truth with sufficient clarity and distinctness, then it is clear that I am behaving *correctly and avoiding error* [italics mine]." Descartes, *Meditations on First Philosophy*, 41.

39 See Schmitt, "What's Wrong with Epistemic Circularity," 380. It appears that most of the sources of our beliefs cannot be justified in a way that is not epistemically circular. I do not agree, of course, with Schmitt's conclusions that such circularity is acceptable, at least in the context we are concerned with in this paper.

40 See for example, Sosa, "Philosophical Skepticism and Epistemic Circularity," and Schmitt, "What's Wrong with Epistemic Circularity."

To conclude, contra van Cleve, his proposed solution of the Cartesian Circle does not shed any new light on the problem of the criterion for epistemology. In the final analysis, the problem of the criterion and its manifestation via the Cartesian Circle loom ever more starkly over the foundationalist landscape. We may exchange logical circularity for epistemic circularity and gain a little credibility, but that is scant consolation against skepticism. The way out of Descartes' Circle must lie elsewhere, if indeed there is one.

Works Cited

- Alston, William P. *The Reliability of Sense Perception*. Ithaca, N.Y.: Cornell University Press, 1993.
- Chisholm, Roderick. *The Foundations of Knowing*. Minneapolis, MN: University of Minnesota Press, 1982.
- Descartes, René. *Oeuvres De Descartes*, 11 vols., eds. Charles Adam and Paul Tannery. Paris: Librairie Philosophique J. Vrin, 1983.
- Descartes, René. *The Philosophical Writings Of Descartes*, 3 vols., translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch, volume 3 including Anthony Kenny. Cambridge: Cambridge University Press, 1988.
- Descartes, René. *Meditations on First Philosophy*. Trans. John Cottingham. Cambridge: Cambridge University Press. 1996.
- Gewirth, Alan. "The Cartesian Circle," *Philosophical Review* 50 (1941): 368-95.
- "The Cartesian Circle Reconsidered," *The Journal of Philosophy* 67 (1970): 668-685.
- "Descartes: Two Disputed Questions," *The Journal of Philosophy* 68 (1971): 288-296.
- Lammenranta, Markus. "Circularity and Stability" Paper presented at *Twentieth World Congress of Philosophy*, Boston, Massachusetts U.S.A. 10-15 August 1998. Retrieved from <http://www.bu.edu/wcp/Papers/TKno/TKnoLamm.htm>

- "Reliabilism and Circularity," *Philosophy and Phenomenological Research* 56(1) (1996): 111-124.
- Loeb, Louis E. "The Cartesian Circle," *The Cambridge Companion to Descartes*, ed. John Cottingham. Cambridge: Cambridge University Press, 1992.
- Newman, Lex, and Alex Nelson. "Circumventing Cartesian Circles," *Noûs* 33(3) (1999): 370-404.
- Reid, Thomas. *Inquiry and Essays*, ed. by R. E. Beanblosson and K. Lehrer. Indianapolis: Hackett Publishing Company, 1983.
- Schmitt, Frederick F. "What is Wrong with Epistemic Circularity?" *Philosophical Issues, Epistemology* 14 (2004): 379-402.
- Sextus Empiricus. *The Skeptic Way: Sextus Empiricus's Outlines of Pyrrhonism*, trans. Benson Mates. Oxford: Oxford University Press, 1996.
- Sosa, Ernest. "Philosophical Skepticism and Epistemic Circularity," *Proceedings of the Aristotelian Society* 68 (1994): 263-90.
- Stroud, Barry. "Understanding Human Knowledge in General," in *Knowledge and Skepticism*, ed. Marjorie Clay and Keith Lehrer. Boulder: Westview, 1989.
- van Cleve, James. "Foundationalism, Epistemic Principles, and the Cartesian Circle," *Philosophical Review* 88 (1979):55-91.

THE ONTOLOGICAL STATUS OF DREAMS IN BERKELEIAN METAPHYSICS

KENNY PEARCE

The University of Pennsylvania

Introduction

The philosophical writings of Bishop George Berkeley have been widely criticized, notably by Immanuel Kant, for their deep reliance on revealed theology and their explicitly stated apologetic purpose. However, beneath the religious polemics and *ad hominem* attacks on theological opponents is a metaphysical system backed by solid, logical argumentation. This system would be not only viable but truly compelling if a few difficulties were to be solved. Two such difficulties arise for Berkeley on the subject of dreams. Firstly, the dream argument Berkeley employs to establish the non-existence of matter seems to contradict the argument used to establish the existence of other human minds, leaving him in a state of solipsism. Secondly, the most fundamental principle of Berkeley's immaterialism - "esse est percipi" - seems to imply that dreams are just as real as waking life, an assertion which Berkeley's commitment to "common sense" renders untenable for him. In answer to these difficulties, this paper will propose a possible metaphysical and phenomenological explanation of dreams based upon the principles of Berkeley's immaterialist philosophy. The first step will be to give an outline of the relevant points from the core of Berkeley's thought and the arguments he employs to get to them. It is in the course of this discussion that the alleged contradiction between Berkeley's dream argument and his argument for the existence of other minds will be dealt with. Next, the dream experience will be examined, and the characteristics of this experience that make it so problematic will be highlighted. In the third section, the general epistemological problem of dreams ("dream skepticism") will be examined, along with its transformation into a metaphysical problem by Berkeley's system. Section four will examine a possible solution to the epistemological problem that was proposed by Leibniz. Finally, in the fifth section, it will be shown that the differences between dreaming and waking perceptions on which Leibniz's epistemic criteria rely allow the Berkeleian to draw an ontological distinction between worlds and the actual world. In particular, an ontology will be proposed which will have

four levels: the level M of minds (or spiritual substances), the level RP of “real” (waking) perceptions, the level DP of dreamed or hallucinated perceptions, and the level T of thoughts. The bulk of that section will be used to demonstrate that Berkeley’s basic view of the “real” world entails that it has certain properties which, according to Leibniz and our common use of language, dreams do not. Because perceptions must be definitive of reality in any Berkeleian ontology, the epistemic criteria proposed by Leibniz, whereby we apply certain tests to our perceptions to determine whether they form part of the “real” world, become a very natural definition of reality and unreality for the Berkeleian. This sort of explanation is necessary because I do not think it is possible for Berkeley to claim that dreams are generated internally by the mind of the dreamer; he must postulate an external source of dreams for the same reason he must postulate an external source of waking perceptions. The characteristics of dreams that make this the case will be discussed in section three.

1. The Berkeleian Position

Perhaps the most fundamental assumption made by Berkeley, and a possible weakness of his theory, is his assertion that “an idea can be like nothing but an idea.”¹ This is essentially an axiom of the system and, while Berkeley takes some pains, in both the *Principles* and the *Dialogues* to show it to be the case with regard to specific sense perceptions, the only general defense of this premise given in Berkeley’s published works is the statement that he, personally, cannot imagine how it could be otherwise.² However an argument is given in Berkeley’s unpublished *Philosophical Commentaries* section 378.³ The premises of the argument are that (1) two things cannot be alike unless they can be compared to one another and found to be similar, but (2) a mind can only compare its own ideas with one another, since they are its only direct objects⁴, and (3) nothing other than a mind could possibly make a comparison between any two objects. The conclusion of the argument is that (4) nothing that is not an idea can be an object for comparison, and, therefore, nothing other than an idea can be like anything.⁵ It follows

1 George Berkeley, *A Treatise on the Principles of Human Knowledge*, 8

2 See, e.g. *loc. cit.*

3 See the discussion at Kenneth P. Winkler, *Berkeley: An Interpretation* (New York: Oxford University Press, 1989) 145- 148.

4 See n. 13 below and Berkeley, *Philosophical Commentaries* 808.

5 To state the argument more formally, define the following predicates:

Lxy =df. x is like y

from this argument (though Berkeley does not explicitly assert this) that it is incoherent to say that a substance is like anything. Only ideas can be like or unlike because it is impossible for comparisons involving non-ideas to be made.

From this principle, it follows that, if matter as a metaphysical entity exists at all, it can in no way resemble our perceptions of it. In light of this, it is difficult to see how our perceptions can be said to be perceptions of matter at all. The course Berkeley anticipates his opponents will take is to argue that all properties of objects are what John Locke refers to as “secondary qualities;” that material objects in no way resemble our perceptions, but what the object really possesses is “only a power to produce those sensations in us: and what is sweet, blue, or warm in idea, is but the certain bulk, figure, and motion of the insensible parts in the bodies themselves.”⁶ However, Locke can assert this only because he thinks that there are other qualities, those he calls “primary,” which do resemble our perceptions. If Berkeley’s principle is accepted and primary qualities are annihilated, then the reason for our belief in matter, or what it is that this hypothesis is supposed to explain, is unclear. Berkeley points out that “it is granted on all hands (and what happens in dreams, phrensies, and the like, puts it beyond dispute) that it is possible we might be affected with all the ideas we have now, though no bodies existed without, resembling them.”⁷ Next, Berkeley anticipates that someone might claim that

Though we might possibly have all our sensations without [material objects], yet perhaps it may be thought easier to conceive and explain the manner of their production, by supposing external bodies in their likeness rather than otherwise; and so it might be at least probable that there are such things as bodies that excite their ideas in our minds.⁸

This, however, does no good, as it only brings up the infamous “mind-

Cxyz =df. x compares y to z

Mx =df. x is a mind

Ix =df. x is an idea

The argument can then be formulated as follows:

$\forall x \forall y (Lxy \rightarrow \exists z Cxyz)$

$\forall x (Mx \rightarrow \forall y \forall z [(\sim Iy \vee \sim Iz) \rightarrow Cxyz])$

$\forall x (\sim Mx \rightarrow \forall y \forall z \sim Cxyz)$

$\therefore \forall x (\sim Ix \rightarrow \forall y \sim Lxy)$

6 John Locke, *An Essay Concerning Human Understanding* II.viii.15, emphasis original.

7 Principles 18

8 *ibid.* 19

body problem” which has vexed philosophers for centuries.⁹ It is, in fact, incredibly difficult to explain the source of our sensations by appeal to physical objects. How is it possible for matter, which is, by definition, inert and non-thinking, to have any effect on a mind whatsoever? Some philosophers would solve the problem by considering the mind as merely a function of the physical brain, so that matter is only affecting other matter. Berkeley’s answer to this strays from pure deductive logic and makes implicit use of Occam’s Razor; that is, Berkeley claims that his immaterialism is a simpler hypothesis with greater explanatory power than the competing pure materialist view.¹⁰

Berkeley’s objection to the existence of matter is that if it exists it is by nature unknowable. This difficulty concerning our knowledge of matter had earlier been acknowledged by Locke who asserted in several places that we have no idea of substance in general, but only postulate its existence for explanatory purposes.¹¹ Berkeley’s dream argument shows that matter is unnecessary to an explanation of the phenomena in question; his argument from the mind-body problem purports to show that it is also unhelpful.

The line of reasoning Berkeley uses to establish his own view is something like the following: I have direct and immediate experience of only one thing: my own mind. My mind has two important properties: it is active (i.e. it wills), and it has ideas. I do not have an idea of my mind, and cannot, since my mind is the thing ideas exist in, and ideas cannot resemble anything that exists outside of a mind. I do, however, have what Berkeley calls a “notion,” which is to say a collection of mostly relational ideas, which give meaning to the word mind.¹² I have specific

ideas, and from these can form a notion of an idea in general.¹³ I also

9 “Though we give the materialists their external bodies, they by their own confession are never the nearer knowing how our ideas are produced: since they own themselves unable to comprehend in what manner body can act upon spirit, or how it is possible it should imprint any idea in the mind,” *loc. cit.*

10 *ibid.* 26

11 See esp. *Essay* II.xiii.1-2

12 This technical usage of the word “notion” was introduced in the second edition of the *Principles*. See esp. sect. 142.

13 The “general notion,” which has the same function in Berkeley’s epistemology as what other philosophers call an “abstract idea,” since Berkeley rejects the concept of abstraction, seems to be a sort of decision procedure by which it can be determined whether a given idea or notion belongs to a class of ideas or notions denoted by a particular word or other symbol. For instance, the general notion of a triangle would be a list of ideas or notions that form a part of any (complex) idea that is to be called a triangle. See esp. the introduction to the *Principles*, sect. 15. Berkeley uses the word “idea” to signify

have ideas of specific instances of the relationship called “inherence” (or “support,” intended in the sense in which substances support accidents in Aristotle), and from this I can form a general notion of inherence. From these two general notions, I can form a further notion; the notion of mind¹⁴ as the substance in which ideas inhere. More information can be added to this notion by similar processes, based on the experience of my own mind that I have. I observe that some of my ideas, those classified as “perceptions,” are not generated by acts of my own will, but are pressed upon me. Accordingly, I conclude that they have a cause outside myself. This cause needs two characteristics: it must be active and it must have ideas. If I were to deny that it was active, then I would not have explained the phenomenon, as I would still need to explain what caused this inert object to do whatever it does to cause my ideas. If I were to deny that it had ideas, then where would the ideas come from?¹⁵ Therefore, Berkeley concludes, my perceptions are caused by another mind.

Anyone who attempts to explain perceptions by means of matter, rather than of an external mind, faces great difficulties. First, we have no other reason to suppose that there is any matter anywhere than that it is allegedly useful in this explanation. By contrast, I already know that at least one mind (my own) exists. Second, matter is by definition inert, so we must explain how it was set in motion in order to ultimately cause my ideas. This is by no means impossible, but adds to the complexity of our hypothesis tremendously. Third, we still have to grapple with the mind-body problem somehow, because we have to get to our subjective experience of perceiving objects from the physical operation of those

objects. This too is possible. The explanation can be made in terms of neurons, chemical activity in the brain, and so forth, but, again, the simplicity of the explanation is lost so that Occam’s Razor recommends that we choose Berkeley’s external mind hypothesis over the matter hypothesis.

The next difficulty Berkeley must face is another that has long plagued philosophy; he must escape from solipsism by providing an epistemological grounding for the belief that minds other than his own

any direct and immediate object of the mind, a convention adopted from Locke (see Locke, *Essay* I.i.8).

14 Berkeley often prefers the term “spirit” which I have not used because its usage appears to be an attempt to insert additional theological doctrines not supported by the argument.

15 One answer, of course, is to suppose that the ideas are innate, already being present in the mind and merely being excited by the presence of the objects. Berkeley, working as he is within a broadly Lockean framework, does not seriously consider this view.

exist. Berkeley's case is somewhat different from many others, as he has already established the existence of one other mind, which he calls God.¹⁶ Berkeley steps lightly over the difficulty in question, noting that "I perceive several motions, changes, and combinations of ideas, that inform me there are certain particular agents like myself, which accompany them, and concur in their production."¹⁷ He later clarifies,

A human spirit or person is not perceived by sense, as not being an idea; when therefore we see the colour, size, figure, and motions of a man, we perceive only certain sensations or ideas excited in our own minds: and these being exhibited to our view in sundry distinct collections, serve to mark out unto us the existence of finite and created spirits like ourselves.¹⁸

The key phrase in this passage is where Berkeley tells us that our perceptions "serve to mark out" the other minds. Early in his life, Berkeley developed the belief that sense perceptions form a language by which the originating mind (God) communicates information to us. This is one of the primary contentions of his *Essay Toward a New Theory of Vision*, the first major philosophical work he published.¹⁹ His claim is that the whole of the physical world is a series of signs, which always have the same meaning. We shall later see that this consistency is a critical feature of reality. For now it suffices to say that we observe the physical world to be consistent; that is, we observe that bundles of sense perceptions always have the same meaning so that when we are mistaken in our predictions based upon them, there is always some error to be found in our reasoning.²⁰

This is where the alleged contradiction in Berkeley's argument takes place. It has been argued²¹ that due to a contradiction between this part of his argument and his earlier dream argument, Berkeley either never gets his system off the ground or is trapped in solipsism. The claim is

16 It has been established that this mind is more powerful than a human mind (as it can force ideas upon human minds directly, which human minds cannot, and it can apparently hold the entirety of the physical world in its memory at once, determining what perceptions to give humans next according to set consistent physical laws) and that it is responsible for sustaining the physical world. It has not, however, been shown to be in any way infinite, or even the originator of the world. As such, whether it resembles the traditional religious conceptions of God closely enough to justify the use of that name is an open question.

17 *Principles* 145

18 *ibid.* 148

19 See especially sect. 147. Also see *Principles* 43.

20 But see subsection 5.5 on mistaken judgments below.

21 Charles Byron Cross, "Berkeley on Other Minds," *Auslegung* 6 (1978): 45-50.

that Berkeley first insists that no collection of sense-perceptions could count as evidence for the existence of matter, because we might be dreaming. He then says that a different set of sense-perceptions serves as evidence for the existence of another human mind. How can this be? Could we not likewise be dreaming the perceptions related to the existence of the other mind?

The critical flaw in an objection like the one above is that it does not take account of the differing contexts of the two arguments.²² Berkeley's dream argument is an objection to the prevailing materialist view, whereas his argument for the existence of other minds takes place within his own system. As we shall see in section five, in any broadly Berkeleian metaphysical theory, the criteria by which we know we are awake and not dreaming must become the criteria which define waking life and these criteria will include the assertion that the sense-perceptions of waking life are "effects or concomitant signs"²³ from which we can derive further information about the world. Thus Berkeley can assert that if we reject his view in favor of a materialist theory then no collection of sense-perceptions will necessarily entail the existence of anything, so that someone who defines reality in terms of matter will be stuck not only in solipsism, but also in absolute skepticism. When we are working within Berkeley's own system, on the other hand, any collection of sense-perceptions that satisfies the conditions for waking life will give us certain information about the world. The nature of these conditions will be one of the primary concerns of the remainder of this paper.

2. The Dream Experience

The dream experience has mystified philosophers, sages, and ordinary people alike for centuries. The troubling question of how we tell the difference, both in the moment, and in memories, between dreams and waking life will be the subject of section three. This current section will focus simply on describing the experience itself and coming to a general understanding of why it is so problematic for philosophy in general and Berkeley in particular. The first thing that must be done here is to define just what a dream is. This appears at first glance to be a very simple question, but has in fact been the subject of much debate. Norman Malcolm, in line with his broader verificationist views, has

²² For more on this subject, see Russell Wahl and Jonathan Westphal, "Descartes, Leibniz and Berkeley on Whether We Can Dream Marks on the Waking State," *Studia Leibnitia* 1992 24(2): 177-181.

²³ *Principles* 145

asserted that a dream just is a dream narration.²⁴ This has a number of problems. First, this essentially assumes that dreams come into existence not during sleep but at the moment of waking. Dream memories are relevantly similar to waking memories, such that if we are to say that they are not generated at any definite time, then we must journey even deeper into skepticism, so that our memories do not necessarily conform even to subjective experiences. Second, Malcolm had little faith in the then-emerging technology of the electroencephalograph (EEG), which now shows great promise in the possibility of actually being able to map different individuals' subjective experiences, such that we might be able to detect the moment of dream generation. Third, Malcolm's view rejects a straightforward, intuitive, common sense view without any persuasive reason for doing so.

More recently, Brian O'Shaughnessy has suggested a more plausible account of the basic nature of dreams.²⁵ O'Shaughnessy asserts that dreaming "is the mind creating out of its own resources an unreal replica of waking consciousness."²⁶ In particular, he enumerates three characteristics of waking life that the dream experience counterfeits: a world in which events take place, mental processes, and perceptual awareness of the world. It is not clear, however, how these things are counterfeits. Dreams really do occur in a world (although not the real world), real mental processes seem to take place in them, and one of these mental processes is perceptual awareness of the world. Can we really be mistaken about the operations of our own conscious (or "as-if conscious" in O'Shaughnessy's terminology) minds? Does it make sense to say that these things do not really occur, that we are deceived

in believing that we were having some subjective mental experience? Certainly there may be cases of false memory, even with regard to subjective experiences, but if our memory is so reliable a tool with regard to our internal mental experiences while awake, then how can we say that this is not the case in sleep?

Of course, dreaming is, by definition, not actual (physiological) consciousness, and this is the point O'Shaughnessy is getting at. The intuitive reasons for this are that dreams take place during sleep (the opposite of physiological consciousness) and the world perceived in dreams is not the same world perceived during physiological consciousness. This is only the beginning of an explanation of the

24 See Norman Malcolm, *Dreaming* (New York: Humanities Press, 1959), esp. ch. 11

25 Brian O'Shaughnessy, "Dreaming," *Inquiry* Dec. 2002, 45(4): 399-432

26 Ibid. 399

differences between dreams and waking life. Any rigorous explanation will require elaborate metaphysical underpinnings, and so will be postponed until the fifth section of this paper.

In dreams, as in waking life, we experience three types of mental objects: perceptions, thoughts, and volitions, where perceptions are ideas entering from the outside, thoughts are ideas of the understanding generated internally by an act of the will, and volitions are those ideas which are the direct objects of the will (as opposed to the understanding). In dreams, however, the distinction between these three categories is blurred. We know experientially that we dream both perceptively and propositionally. That is, we not only dream of having certain sense perceptions (i.e. dream ideas of the sort that might possibly be actual sense perceptions in waking life), but we also dream that certain propositions are true. For instance, in the course of preparing this paper I had a dream in which I was in my mother's house, but the house that I was in was most certainly not the one in which my mother lives. That is, I dreamed that the proposition expressed by the sentence "I am in my mother's house" was true, while simultaneously dreaming a bundle of sense perceptions which, were I awake, would count as conclusive evidence against the truth of that proposition. There is nothing unusual about this sort of dream, as dreams are most certainly not limited by any contingent facts of the actual world.²⁷ A particularly interesting point about this sort of dream is that within it the proposition expressed by "I am in my mother's house" is actually a perception rather than a thought. It is an idea that comes in from the outside, which is an immediate and inescapable "truth" of the dream world. I do not reason about my sense perceptions to come to this conclusion, but simply know its truth, in the

same way that I know the truth of a statement such as "the door is brown." Furthermore, it seems to me that there are some dreams in which we perceive our own volitions, which is to say that ideas that would ordinarily be volitions, products of our will, are in fact imposed on us from without by the dream world. Likewise, the blurring sometimes goes the other direction. There are cases in which what would ordinarily be perceptions are, in the dream world, volitions.

What these observations point out is the wild variation in the degree of control we can exercise over our dream worlds. In some cases, everything in our dream world, including our own actions, is imposed from the outside and we have no control over it. In other cases, a simple

²⁷ Whether dreams are limited even by the laws of logic is a subject of debate. On this, see below and also O'Shaughnessy, "Dreaming" 409-410.

act of the will is able to drastically alter the world around us.

The above discussion makes clear one of the most serious problems with dreams for the Berkeleian. In dreams we do experience perceptions. There are certain ideas in dreams which we must acknowledge have, as was said above, "been imposed from without." They cannot be caused by the mind,²⁸ because they are not products of the will. The materialist can claim that they are caused by the purely physical functionings of the brain (which he may refer to as "the unconscious mind"), but the Berkeleian can obviously not claim this, as the physical functionings of the brain are merely so many sense perceptions of the individual observing the EEG. Rather, the Berkeleian must admit that they are caused by some external mind, just as are our waking perceptions. The obvious course to take is to say that they are caused by the same mind as our waking perceptions, and that this mind is operating in accordance with the physical laws it has devised so that from a physical (rather than metaphysical) perspective, we can indeed explain our dreams in terms of the physical functioning of the brain, while it is understood that the metaphysical reality is that dreams, like the rest of our perceptions, are caused by another mind. We will return to this discussion in greater detail in the section five.

Another important question about the dream phenomenon is the limit of "dreamability." Is there anything one cannot dream? Clearly one can dream things that are false. O'Shaughnessy has suggested that one can also dream the logically impossible,²⁹ but it is not clear that he is correct. Certainly one can dream various propositions simultaneously, and these propositions, taken together, may entail some contradiction, but to say

that "one can dream that 1 and 1 make 3, [or] that one is looking point blank at a surface which is red and blue all over"³⁰ stretches the limits. Perhaps O'Shaughnessy is led into this belief by his subsequent assertion that one cannot dream of having a belief that is false within the context of the dream. I must disagree with this claim as I find that I sometimes have dreams in which I occupy two different perspectives in the dream world: the ordinary limited perspective of myself, and the perspective of a third-person omniscient narrator. In such a situation I have dreamed that, for instance, I was being chased and was not aware of it. Once dream-belief is separated from dream-truth it becomes clear

²⁸ Where the mind is "a thinking thing," which is the normal usage of the word in early modern philosophy.

²⁹ p. 409

³⁰ loc.cit.

that dreaming that “I heard the sentence ‘1 and 1 make 3’ and believed it” and dreaming that the proposition expressed by $1+1=3$ was true are two different things, and the latter is probably not possible, even if there is such a proposition.

O’Shaughnessy is, however, correct in asserting that narratability is a critical characteristics of dreams; we cannot dream anything which we could not later narrate coherently. That is, one could probably dream of an invisible pink unicorn (in this case one would actually be dreaming the counter-factual conditional expressed by “if I could see the unicorn, it would be pink [but I can’t, since it’s invisible]”), but one could probably not dream Noam Chomsky’s infamous sentence, “colorless green ideas sleep furiously,” since the latter is so incoherent as to violate the narratability condition, and certainly does not express a proposition. (Although one might be able to dream the proposition “the sentence ‘colorless green ideas sleep furiously’ expresses a proposition, and that proposition is true,” but I would suggest that in this case one would in fact be dreaming that the sentence had syntactic structure and/or semantic content other than what it in fact has in English, and thus in an important sense would be dreaming up a new language. Of course, it is not necessary that one actually know the meaning of the language in question; the dream could merely claim that there was such a language without that language having any definite interpretation.)

To summarize, the dream experience is one very much like that of waking life, but conventional wisdom insists that dreams do not give information about the “real world.” It is not clear how the “real world” is to be defined, but clearly “the world we experience while awake” will not do. How would we then define wakefulness? Dreams involve every type of idea the mind is capable of experiencing, but these ideas may not function in the same way or have the same origin (internal or external to the dreamer) as they do in waking life. All of these characteristics lead to major philosophical problems, which will be the subject of the section to follow.

3. Dream Skepticism and the Associated Difficulties for Berkeleian Metaphysics

The primary problem of dreams, the one that has plagued philosophers of all stripes, is epistemological. How can I be certain that I am awake and not sleeping? If I can’t, how can I be sure of anything? This

problem is universal and deeply troubling. More troubling still are the metaphysical problems it creates for the Berkeleian and the Idealist/phenomenalist School of philosophy generally.

Part of what Wahl and Westphal say³¹ in defense of the maneuvers taken by Descartes, Leibniz, and Berkeley to escape dream skepticism is that in dreams we do not actually reason but only dream of reasoning. Malcolm takes this even farther, to assert that dreamed experiences do not occur at all, even subjectively, so that we can always know we are not dreaming.³² If, as I have argued, this is false, the problem runs deeper. From what has been said thus far, no obvious distinction can be made without begging the question. If dreams are distinguished from reality only in that they take place while we are asleep, and they involve very real mental activity, then how can we ever know that we are awake?

In Berkeleianism in particular, the problem becomes even more disturbing, due to the basic principle that “to be is to be perceived.” If this is the case, where is the Berkeleian to place dream perceptions in his ontology, and how can he avoid asserting that dream objects actually exist? Another idealist, Arthur Schopenhauer, faced the same problem, and, rather than making any attempt to overcome it, simply conceded dreams to have the same ontological status as waking life, being distinguished from it only by length.³³ For the Berkeleian, who is determined to “think with the learned, and speak with the vulgar,”³⁴ this answer is unacceptable. He is committed to the development of a theory of reality on which waking life is “real” and dreams are not; or, at the very least, on which dreams are less “real” than waking life.

Berkeley directly considers this problem only once in his major philosophical writings, at *Dialogues* 235, and dismisses it almost instantly:

The ideas formed by the imagination are faint and indistinct; they have besides an entire dependence on the will. But the ideas perceived by sense, that is, real things, are most vivid and clear, and being imprinted on the mind by a spirit distinct from us, have not a like dependence on our will. There is therefore no danger of confounding these with the foregoing: and there is as

31 p. 178

32 See esp. ch. 18

33 Arthur Schopenhauer, *The World as Will and Representation*, E. F. J. Payne, tr. (New York: Dover Publications, 1969), sect. 5

34 Berkeley, *Principles* 51

little of confounding them with the visions of a dream, which are dim, irregular, and confused. And though they should happen to be never so lively and natural, yet by their not being connected, and of a piece with the preceding and subsequent transactions of our lives, they might easily be distinguished from realities. In short, by whatever method you distinguish things from chimeras on your own scheme, the same, it is evident, will hold also upon mine. For it must be, I presume, by some perceived difference, and I am not for depriving you of any one thing that you perceive.

Berkeley's response is so obvious, and so characteristic of his writings, as to make it almost laughable. Philonous is responding to precisely the objection above, that Berkeley's immaterialism is unable to distinguish dreams from waking life, and the content of his response is simply that there must be a perceived difference between the two; if there weren't, the inability to distinguish between them would not be a problem. You perceive this experience to be a dream, therefore it is a dream. "Esse est percipii." We are sometimes mistaken about dreams, but the problem of apparently incorrect perceptions is one that Berkeley deals with at length, so this is not any new difficulty. However, there remains a critical difference between uncertainty about dreams and mistakes in other perceptions. While we do seem, in actual practice, to nearly always know correctly and with great certainty whether a given experience took place while we were dreaming or awake, it is not clear how we perceive a difference. All Berkeley has claimed to do is to reduce the problem of dreams in his system to the point of identity with the problem faced by other philosophers, but, while he does seem to have shown that any solution to the general problem will also solve his particular problem, the difficulty is nevertheless much more urgent for him: if the problem remains unsolved, materialists will merely have lost the ability to distinguish with certainty between dreams and waking life, but for the Berkeleian there will be no such distinction. The difficulty is not merely epistemological, but is a fundamental problem of metaphysics. Furthermore, what of other minds? If this distinction collapses, the Berkeleian will have to concede that either there is ground for thinking that the people we see in our dreams are real people with minds like our own, or else that there is no ground for thinking that other human minds exist at all.

Every one of Berkeley's discussions of "wrong" perceptions centers around the idea that it is not the perception that is mistaken, but our interpretation of it. The information given to us by means of the

language of sense perception was all true, but we failed to understand it. It is not clear, however, that this “language” contains any certain indications as to whether we are dreaming or awake, while at the same time it is necessary to interpret its symbols differently depending on which state we are in. Berkeley’s own attempt to solve this problem, which appeals to the vividness of waking life and the ever-popular Cartesian (external) coherence criterion, is probably not sufficient to provide us with the kind of certainty needed to claim that dreams are less real than our waking experiences, and that the people we meet while awake are real while the people we meet in dreams are not. A stronger set of criteria is needed.

4. The Leibnizian Solution

4.1 *The Cartesian Coherence Criterion.*

Because of the universality of the problem of dreams, almost every major philosopher has dealt it with. The particular treatment of the subject we will focus on in this section is that given by Leibniz in his “On the Method of Distinguishing Real From Imaginary Phenomena.” However, the first criterion he discusses was earlier proposed by Descartes in his Sixth Meditation, and it is in this context that it is best dealt with. Here, in concluding his work, Descartes finally deals with dream skepticism, the doubt that began his series of meditations. “I now notice that there is considerable difference between [dreams and waking life];” he insists, “dreams are never linked by memory with all the other actions of life, as waking experiences are.”³⁵ This external coherence criterion has since been adopted by innumerable philosophers, and it is easy to see why. When asked how we know that something occurred in a dream, one of our most frequent answers is that we remember waking up afterward, and this is enough in and of itself for this test to be of the utmost significance in any philosophical discussion of dreams. However, this is far from sufficient for the kind of certainty that is required to succeed in escaping from dream skepticism.

The obvious objection, which has been given by philosophers since this criterion was first proposed, is that there is nothing to prevent someone from dreaming that this test was satisfied. Descartes’ answer is that if the test were satisfied in a dream, then God would be a deceiver. Most philosophers have found this argument unconvincing. For instance, Leibniz summarily dismisses the argument, stating that “no one will

³⁵ Rene Descartes, “Meditations on First Philosophy” 89 in John Cottingham, Robert Stoothoff, and Dugald Murdoch, trs. *Descartes: Selected Writings* (New York: Cambridge University Press, 1988), 122

fail to see how far it is from a demonstration having metaphysical certainty, for we are not deceived by God but by our judgment, asserting something without accurate proof."³⁶ Furthermore, Nicolas Malebranche suggests a story³⁷ in which a man on several different occasions goes to sleep at night and dreams that he has risen from bed, gone out to a "black Sabbath," and practiced sorcery. Because the events of the dream always occur at night, and always end with the dreamer returning to bed before morning, the external coherence criterion is no help, and he continues to believe that his experiences really occurred. Less outlandishly, one individual with whom I discussed the subject of this paper reported that from time to time he has dreamed the occurrence of various mundane events, especially conversations with friends, and later been surprised when the friends in question do not remember the conversation, or he otherwise discovers the event not to have actually occurred.³⁸ These sorts of dreams are not unlike the dreams we all experience regularly in any relevant way. This counts as clear evidence that the Cartesian coherence criterion alone is insufficient for the task at hand. Recognizing the deficiency of Descartes' proposal, Leibniz gives three additional tests, which are different from the one discussed above in that they deal only with the internal content of the dream, and require no additional knowledge of the outside world.

4.2 Vividness

The first of Leibniz's internal criteria is vividness. It will be recalled that this was also one of the tests Berkeley mentioned. Clearly this test is not sufficient on its own. First, everyone has had dreams vivid enough that he believed the dream to be real (at least until he woke up at the end). Second, as Schopenhauer observes, "no one has ever held [dreams and waking life] up to comparison; only the recollection of the dream could be compared with the present reality."³⁹ This comparison is clearly unfair. Our recollections of past waking life are not always vivid either; how can we be certain that the dream was not just as vivid as waking life at the moment it occurred? Nevertheless, the lack of vividness in dream memories is, in practice, a criterion we use to distinguish them from memories of waking life. Even if this is only because the disconnect between the dream and the moment at which it is being examined (i.e.

36 G. W. Leibniz, "On the Method of Distinguishing Real From Imaginary Phenomena," in *Philosophical Papers and Letters*, ed. Leroy F. Loemaker (Boston: D. Reidel Publishing Company, 1956): 364.

37 *The Search After Truth* 2.3.6

38 Note that at this point new evidence has caused an experience, which previously fulfilled the Cartesian coherence criterion to no longer fulfill it, so that it is still this criterion that has informed him that he was mistaken.

39 Schopenhauer, *Will and Representation* 16

the experience of waking up) causes the memory to lose its vividness, this still provides an important test for determining at least after the fact the proper interpretation of our sense experiences and, as we shall see later in this section, this is all that is needed in order for Berkeley's general explanation of mistaken perceptions to apply.

4.3 *Complexity*

The second internal criterion Leibniz proposes is complexity. In defining this criterion and distinguishing it from vividness, he says the following:

The phenomenon ... is vivid if its qualities, such as light, color, and warmth, appear intense enough. It will be complex if these qualities are varied and support us in undertaking many experiments and new observations; for example, if we experience in a phenomenon not merely colors but also sounds, odors, and qualities of taste and touch, and this both in the phenomenon as a whole and in its various parts which we can further treat according to causes. Such a long chain of observations is usually begun by design and selectively and usually occurs neither in dreams nor in those imaginings which memory or fantasy present, in which the image is mostly vague and disappears while we are examining it.⁴⁰

In dreams, we rarely (or, more likely, never) undertake examinations of the depth Leibniz suggests, but his point is not that this in itself, the experience of examining the phenomenon, proves that it is not a dream. Rather, the suggestion is that if we consider (after the fact) our dreams, we quickly come to believe that if we had undertaken such a task, the dream would have quickly devolved into nonsense. This is essentially the same as saying that, unlike in waking life, what is not explicitly included in the dream is no part of it, and as such has no truth-value within the context of the dream.⁴¹ Upon examination, dreams are found to be missing an enormous number of details we would think important in the narration of a waking experience, and it is not merely the case that we did not take note of them; they are altogether absent.

4.4 *Internal Coherence*

The final test that Leibniz suggests is the internal coherence criterion. The idea here is that real phenomena operate according to what Leibniz would later call "the principle of sufficient reason." That is, for any real event, there must be a reason why it occurred. This is not the case in dreams. In considering differences in our presumptions with regard to dream narrations and novels, O'Shaughnessy imagines the following

⁴⁰ Leibniz, "On the Method" 363-364

⁴¹ On this, see O'Shaughnessy, "Dreaming" 401-404

scenario:

I dream of copper, of nitric acid, of an encounter between the two, and of nitrogen peroxide coming from the solution. Then whereas concerning a novel we can say 'Andre Bolkonsky died from wounds received at the Battle of Borodino', irrespective of whether it is explicitly stated that he did, we cannot say that 'in my dream the gas was caused by the encounter between the two reagents' – unless it is an explicit part of the dream. And even if it were, I might very well have embellished the report with the further detail that the gas was caused by the encounter of those two reagents only because it was Sunday.⁴²

Similarly to the limitations on complexity discussed above, if these details are not explicitly included in the dream (e.g. if one does not dream a reason for the reaction) then there simply is no reason. We have no expectation of coherence or predictability in our dreams, as we do in waking life, and this is in practice perhaps the most important criterion for distinguishing dreams from reality: dreams, for the most part, fail to make sense.

5. Leibniz's Criteria as a Solution to the Metaphysical Difficulty

Leibniz is not particularly satisfied with the criteria offered above. In his own words,

We must admit it to be true that the criteria for real phenomena thus far offered, even when taken together, are not demonstrative, even though they have the greatest probability ... Thus by no argument can it be demonstrated absolutely that bodies exist, nor is there anything to prevent certain well-ordered dreams from being the objects of our mind, which we judge to be true and which, because of their accord with each other, are equivalent to truth so far as practice is concerned.⁴³

Whatever his mature view may be, in this passage, which comes from a very early work, Leibniz seems to be speaking as a materialist,⁴⁴ and as such is prone to skepticism. He does suggest that the objects of our perceptions, if sufficiently well ordered, may be "equivalent to truth so far as practice is concerned," but implicit here is the idea that they are not actually truth for theoretical, rather than practical, purposes unless they correspond to mind-independent physical objects. Because he here seems to link the veracity of our perceptions to accurate representation

⁴² Ibid. 403

⁴³ Leibniz, "On the Method" 364

⁴⁴ In Berkeley's sense: that is, as one who believes in matter as a metaphysical entity, not as one who denies the existence of the immaterial.

of external physical bodies, he does not seem to be in a position to banish dream skepticism completely. The Berkeleian is now in a much stronger position, and once again able to take advantage of the greatest virtues of his system: namely, its simplicity, and its compatibility with “common sense.”

Because the Berkeleian asserts that “to be is to be perceived,” once we have demonstrated that there exist criteria for differentiating dreams from waking life, and it can be seen that we do in fact use these criteria, he is in a position to differentiate dreams from reality ontologically. After all, he is “not for depriving you of any one thing that you perceive,” and if you do in fact apply the above criteria, then you perceive a difference between dreams and waking life. The Berkeleian position is further strengthened by the exact nature of these criteria, in that, given various other basic principles of Berkeley’s philosophy, particularly the view of sense perception as a language, it is perfectly natural that some of them, in particular the internal coherence criterion, should be used to define reality.

5.1 A Berkeleian Ontology

The Berkeleian will want an ontology with at least four basic levels: The level M of minds, the level RP of “real” perceptions, the level DP of dreamed or hallucinated perceptions, and the level T of thoughts and volitions. There may be differentiation within these levels. For instance, there may be reason to suppose that the mind, which is the source of our perceptions, is on a higher ontological level than we are. However, there will most likely not be so sharp a divide between this mind and our own as there is between our minds and their ideas.

The divide between M and the other levels is a simple and straightforward consequence of Berkeley’s overall theory. Ideas are real (i.e. they occupy some level in our ontology), but they exist only within minds, and are dependent upon minds. Therefore, they are “less real” than minds. What is not so clear is why ideas are to be subdivided into the three levels mentioned. What relevant features of RP distinguish it from DP, or even T?

The entities at level T are unique in that they are dependent only upon the will of the mind in which they reside. They are not imposed on that mind from without, nor does that mind impose them upon or share them with others. They also lack the vividness of perceptions. These characteristics all seem relevant, and sufficient to distinguish them from perceptions. The division between RP and DP is more difficult.

The four criteria listed by Leibniz are all used by human beings to distinguish between the two in practice, but is any of them relevant to the ontological problem? In fact, all of these criteria lead to important differences between entities at RP and entities at DP, and can be used to ascribe to RP several qualities essential to the conception of physical reality implicit in Berkeley's writings.

5.2 The External Coherence of Entities at RP

Entities existing at level RP, that is, perceptions of the "real" physical world, satisfy the Cartesian coherence criterion. They form part of a long chain of perceptions that fit together to form a unified picture of a world with a history and consistent laws of cause and effect. If this were not the case, if our perceptions grouped themselves into discrete incidents, each internally consistent but not connecting to the others in any way, we could never learn the "language" of the physical world, and so we could not conclude anything from our perceptions. One long sense experience with room for experimentation is needed in order for this communication to take place. Dreams may be meaningful, that is, they may communicate true information to us, but because they are so short we cannot derive the language of a dream from the dream itself. This is why every scheme of dream interpretation throughout all of history has somehow incorporated information external to the individual dream. Today, while it is no longer popular to believe that dreams predict the future, many people believe that dreams are of psychological importance, that they reveal deep truths about the mental state of the dreamer. This sort of dream interpretation relies on the meaning of the dream content to the dreamer himself, meaning he has attached to these symbols while awake. Other types of dream interpretation have relied upon standardized types of symbols, which are thought to be applicable to dreams in general, independent of their meaning for the individual dreamers. Still others, particularly the Hebrew prophets, have claimed to have dreams or visions that ended in an encounter with a figure that interpreted the foregoing dream content in normal spoken language. These dreams have a limited degree of external coherence which assigns to them some limited degree of meaningfulness. If a sense experience were completely isolated, having no external coherence whatsoever, we would be unable to assign any significance to it. Thus the connectedness of one period of wakefulness to another is what makes waking perceptions meaningful, and, since one dream is generally not connected with another, dream perceptions become meaningful only insofar as they too share some limited connection with the series of perceptions making up waking life.

5.3 *The Vividness and Complexity of Entities at RP*

The Berkeleian wants very much to assign responsibility for the generation of our perceptions to a mind much greater than our own. In particular, Berkeley himself wants to say that the source of our perceptions is the God of orthodox Christianity. It is not at all clear that Berkeley's arguments get us nearly so far as this, but he has certainly provided sound footing for a theory of natural theology. Every perception we have reveals to us information about the mind that is its source. The external coherence of our waking perceptions, and also Occam's Razor, makes it rational for us to assume that all of our waking perceptions are caused by a single mind. Their vividness is indicative of the additional powers that mind possesses beyond our own, and their complexity is indicative of that mind's intelligence. Vividness is indicative of the power of the divine mind, in that not only is that mind able to place ideas in our minds, but the ideas it places there are actually stronger than the ideas we create ourselves. The complexity of our waking perceptions are indicative of the intelligence of the divine mind in that this mind is able to hold the entirety of the physical world within it, remembering every detail and being able to determine on the basis of these details what perceptions to give next.

What has been said in this section leads to a startling conclusion. Dreams may not, in fact, have the divine mind as their origin at all. Dreams are so small, so simple, so incoherent, and so often copied from waking experiences that they need not be caused by a mind very much greater than our own. All a mind needs to cause dreams is the ability to impress perceptions on other minds, and that vaguely. Our own level of creativity and intelligence will suffice. However, we here run afoul of Occam's Razor again. Remember that, for all the same reasons that waking perceptions cannot have their origins in our own minds, dreaming perceptions must also have an external source. If a Berkeleian had independent reason to believe in the existence of minds comparable in intelligence to our own but having the power to impose their ideas on others,⁴⁵ it would be perfectly natural for him to suppose that these were the cause of his dreams, and not the mind, which causes his waking perceptions. However, in the absence of such independent ground, it is far better for the Berkeleian to stick close to the orthodox

⁴⁵ I have in mind angels or other "spiritual" beings as possible sources of dreams. In fact, theories stating that dreams are caused by some sort of angel or minor deity are not uncommon in the history of thought about dreams. See Ian Hacking, *Historical Ontology* (Cambridge, Massachusetts: Harvard University Press, 2002) 235-236, 254, and also *Iliad* 2.1-34 where Zeus addresses an (apparently sentient) dream and commands it (him?) to deceive Agamemnon. Some editors of the Greek text even capitalize the word for dream, treating it as a proper noun.

scientific viewpoint that dreams can be explained in psychological terms and claim that the divine mind causes them in accordance with the set laws all of our other perceptions obey. This has the added advantage that it successfully fits dreams into the same world as the rest of our perceptions.

5.4 *The Internal Coherence of Entities at RP*

Our waking perceptions make sense. They obey predictable laws. As Leibniz tells us, in waking life we are never deceived when we make good use of reason. If a false conclusion is formed from our waking sense perceptions, a reasoning mistake has been made. This is for the Berkeleian, even more so than for Leibniz, the critical feature of waking life. To say that a set of perceptions has internal coherence is simply to say, in Berkeley's terminology, that they form a single language with a single correct interpretation. Not only does this property get the Berkeleian the general view of sense perception he wants; it is also critical to establishing the existence of other human minds. The argument might be given as follows:

- 1) The entities at level RP are (by definition) the vocabulary items of a language
- 2) The vocabulary items of a language must have a consistent interpretation in order for successful communication to occur.⁴⁶ If no successful communication occurs, the system in question is not a language.
- ∴ 3) The entities at level RP have a consistent interpretation
- 4) There exists a set B of entities at level RP (namely, my body) which is properly interpreted as communicating information about my mind
- 5) There exist other sets of entities at level RP that are relevantly similar to B (i.e. other human bodies)
- 6) If two sets of vocabulary items are relevantly similar to one another, they will have relevantly similar interpretations
- ∴ 7) There exist other minds relevantly similar to my own

The first premise of this argument can be objected to on the ground that, while it is clearly legitimate to define RP in this way, there may not be a collection of perceptions meeting this requirement. That is, it may follow that nothing is, in the relevant sense, "real." I do not think this objection

⁴⁶ Lexical ambiguity notwithstanding. Lexical ambiguity dampens effective communication, such that when the ambiguity goes unresolved the intended piece of information is not communicated. As a result, it is conceivable that there could be a system of sounds or visible symbols intended as a language that had so much unresolvable lexical ambiguity that it might not truly qualify as a language at all.

is particularly important. What it points out is that a consequence of this sort of ontology is that if none of our sense perceptions give us information about the world, then there is no physical world. For any philosopher of the idealist/phenomenalist persuasion, this should be a perfectly acceptable conclusion. After all, any materialist should concede that if an individual's sense perceptions do not communicate any information about the world to him then he is not connected to physical reality. Since "to be is to be perceived," a physical reality disconnected from minds is no reality at all. Besides this, it is an everyday experience that we are able to apply reason on the basis of standard interpretations of our sense perceptions to come to correct conclusions about the world, and in particular about what sense perceptions we will experience next.

The second obvious objection is that the phrase "relevantly similar" in (5) through (7) is doing a lot of work. It is not clear what "relevantly similar" means in these differing contexts (the context of a body and the context of a mind). There is an extent to which this objection is valid. We are arguing from a relevantly similar body to a relevantly similar mind, when the mind in no way resembles the body and, in fact, cannot resemble anything.⁴⁷ Indeed if the body were a purely static entity this objection would probably be inescapable; but if the body were purely static we would say it was dead and not attempt to claim that there was any mind associated with it. Instead, I observe that other human bodies look enough like my own for me to conclude that they are the same sort of thing (an object which might conceivably be the instrument of my mind or of another like it), and that they are dynamic. They act. In observing the actions of other bodies, I notice that they are similar to my own in that they seem to act intelligently based upon definite goals, and they also speak intelligently, producing at least the illusion of communication. In the case of my own body, I know that all of these things occur only when I will them to, and I know that a rational thought process often precedes these volitions. Thus I may correctly conclude that these other bodies are controlled by minds that are similar to my own in various ways corresponding to the similarities in the actions undertaken by the bodies.⁴⁸ Since the external and internal coherence of our total waking experience, features not shared by dreams, permit waking perceptions to function as a language while dream perceptions cannot (except insofar as they are connected with waking life), it is

⁴⁷ See sect. 1 above.

⁴⁸ *But how can they be "similar" to my mind if they are not "alike" and do not "resemble" each other? What is the difference between these words? I believe that a response can be given by resorting to Berkeley's theory of "notions" and claiming that we are able to form notions of minds and notions, being ideas, can be compared to one another, but we do not truly have even a notion of matter.*

correct to conclude that the human bodies we see while awake are animated by human minds, while those we see while dreaming are not.

5.5 *Mistaken Judgments*

An unfortunate problem with all of this, a difficulty the Berkeleian must face constantly, is the fact of mistaken perceptions. I do not believe I've ever spoken with an individual who has believed himself to be dreaming when he is in fact awake, but we frequently are certain we are awake while we are dreaming, and I have already mentioned the case of an individual who has sometimes confused dream memories with waking memories. The solution to the Berkeleian metaphysical problem given above relies heavily on our ability to judge correctly whether we are awake or sleeping. What about these mistakes in our judgment?

At *Dialogues* 238, we have the following exchange:

HYLAS. What say you to this? Since, according to you, men judge of the reality of things by their senses, how can a man be mistaken in thinking the moon a plain lucid surface, about a foot in diameter; or a square tower, seen at a distance, round; or an oar, with one end in the water, crooked?

PHILONOUS. He is not mistaken with regard to the ideas he actually perceives; but in the inferences he makes from his present perceptions. Thus in the case of the oar, what he immediately perceives by sight is certainly crooked; and so far he is in the right. But if he thence conclude, that upon taking the oar out of the water he shall perceive the same crookedness; or that it would affect his touch, as crooked things are wont to do: in that he is mistaken. In like manner, if he shall conclude from what he perceives in one station, that in case he advances toward the moon or tower, he should still be affected with the like ideas, he is mistaken. But his mistake lies not in what he perceives immediately and at present (it being a manifest contradiction to suppose he should err in respect of that) but in the wrong judgment he makes concerning the ideas he apprehends to be connected with those immediately perceived: or concerning the ideas that, from what he perceives at present, he imagines would be perceived in other circumstances.

The above considerations have reduced the problem of mistaken judgments about dreams to be the same problem as mistaken judgments about oars. Because there are definite perceived criteria that must be satisfied by waking life, the Berkeleian can now tell the confused dreamer that he was correct in thinking he saw a purple dragon, but his judgment that this purple dragon would form a coherent part of his life, connecting with his other perceptions, was flawed. When there is any confusion at all, it is generally because the four tests outlined generate

mixed results, and the external coherence criterion will sometimes change results over time as new events occur which the experience must fit with in order to be judged real. There may also be times when judgment ought to be withheld, but this, again, is no different than with other types of perceptions and so does not undermine Berkeley's thought to any significant degree.

6. Conclusion

The philosophy of George Berkeley is riddled with simple answers to complex questions, and this is one of the greatest strengths of his system. However, these answers are often vague and incomplete and in his focus on simply establishing the core of his immaterialist thought Berkeley is often content to show that an apparent difficulty in his system reduces to a problem also faced by the materialist, rather than actually proposing a solution. The case of dreams is particularly troubling in that it appears at first glance to bring Berkeley's philosophy to innumerable absurd conclusions. Despite this, Berkeley brushes over it with ease, as though it were nothing. His treatment of the subject is unfortunate in its brevity and lack of detail, but, like most of Berkeley's simple answers, his treatment of dreams at *Dialogues* 235 has the potential to be expanded into a rigorous defense of the immaterialist position. Much more is at stake in the question of dreams for the Berkeleian than the materialist; if the Berkeleian fails to solve the epistemological problem, he is trapped in the assertion that dreams are just as real as waking experiences. I have argued in this paper that not only is there a solution easily available to the Berkeleian, but this solution provides one of many possible examples of how Berkeleian philosophy, for all its simplicity, is essentially stronger than the prevailing views.⁴⁹

⁴⁹ This paper was prepared during the 2004-2005 school year as part of an undergraduate research fellowship from the Penn Humanities Forum, under the supervision of Professor Karen Detlefsen, for which the author is grateful.

SCANLON'S AGGREGATION PROBLEM

TIM CAMPELL

University of Buffalo

1. Introductory Sketch of the Aggregation Problem

According to T. M. Scanlon's contractualist formula, an act is right if and only if it can be justified to others based on principles which no individual would reasonably reject.¹ Justifiability to others provides not only the normative basis of the morality of right and wrong but also identifies the property of rightness itself. The basic content of morality lies in justifiability to others based on their personal reasons to accept or reject certain principles. The structure of Scanlon's contractualism allows only the strongest reasons of a single individual to determine which principles are accepted or rejected. This individualist restriction prevents aggregative moral reasoning. It means that individuals' personal reasons for rejecting principles cannot be summed. Thus, any judgment about the rightness or wrongness of an act can be seen as a comparison of the two strongest individual claims in favor of or against that act. Whoever has the strongest personal reasons for or against the principle will determine whether we should accept or reject it.

Scanlon believes that allowing aggregative moral reasoning would invite reliance on states of affairs as the guiding consideration for our moral thinking. But according to Scanlon, using states of affairs to decide whether or not to engage in a contractualist framework would undermine the basic motivation for his position. A distinctively contractualist position would have to take the idea of justifiability to others as basic, rather than as just the motivation for promoting the best states of affairs.² Scanlon also believes that disallowing aggregative moral reasoning lets contractualism solve certain moral problems in an intuitively appealing way.³ For example, contractualism would not allow the severe burdens of one or just a few people to be outweighed by the small benefits of very many others. In Scanlon's words: "It allows the intuitively compelling complaints of those who are severely burdened to be heard, while, on the other side, the sum of the smaller benefits to others has no justificatory weight, since there is no individual who enjoys these benefits."⁴ We should not feed the Christians to the lions, no matter how many Romans are cheering.

The problem, however, is that there are some cases where numbers do warrant moral consideration. When two different groups of people face the same harm and an agent can save only one of the two groups (at little or no risk to her own safety), it seems obvious that she should save the larger group. However, if in deciding whether a certain course of action is the right one, we are only allowed to consider the strongest personal claims of individuals, then according to contractualism there would be no moral difference between saving one or the other group since each individual's personal claims to be saved are equal. Scanlon believes that this cannot be right. He gives an argument, which I discuss below, that allows for at least the appearance of limited aggregation in such cases. This argument is widely rejected for various reasons which I discuss later. In fact many have argued that since contractualism cannot deal with cases where numbers matter, it should be rejected based on this crucial deficiency.⁵

In this paper I will show that, contrary to what many believe, contractualism is able to solve moral dilemmas involving numbers in an intuitively appealing way. Rejecting Scanlon's proposed solution to those dilemmas does not give us sufficient grounds for rejecting his theory. I will argue that the two most significant problems with Scanlon's argument are (1) that it discounts the importance of burdens that are non-substantive (i.e. burdens that have no actual effect on a person's well being, but which a person would want to avoid; for example, the chance of facing a substantive burden that would affect her well being) and (2) that it takes the frequency of the actual occurrence of certain hypothetical cases involving numbers into account in deciding how to solve those cases within a contractualist framework. These features, far from making Scanlon's argument stronger, place restrictions on the kinds of moral dilemmas involving numbers that Scanlon is able to solve. If Scanlon dropped or altered these features, numbers would no longer pose a problem for his theory.

Here is the layout of the paper: In sections 2 and 3, I will introduce Scanlon's view on aggregation, highlighting some of its problematic aspects. In sections 4-6 I will discuss some very forceful objections to Scanlon's aggregation argument which also suggest that contractualism should be rejected based on the failure of his argument. In sections 7-9, I will show how Scanlon could, in abandoning a few features of his original argument, develop an intuitively appealing solution to problems involving numbers while remaining consistent with the basic idea of contractualism. I'll wrap things up in section 10.

Part I. What Problems Do Numbers Raise for Contractualism?

2. Scanlon's Argument and Its Ability to Deal with Numbers

What does contractualism have to say about cases where numbers seem to matter? Scanlon's suggestion is that we follow what he calls the tie-breaker principle. According to this principle, when the harm that two different groups of people face is the same, numbers break ties, and we ought to save the larger group. The rationale behind this is that if an agent were to save the smaller group, any person in the larger group could complain that the principle which allowed this did not take into account the value of saving her, since it lets the agent decide who to save in the same way that it would have allowed had the numbers of people in each group been equal. Since an individual in the larger group would have this reason, but an individual in the smaller group would not, the former would have the stronger claim to be saved, and thus, an agent would be morally required to save this group.

The tie breaker principle has been challenged by Derek Parfit.⁶ Parfit claims that if an agent were to follow the tie-breaker principle, saving the larger group and ignoring the smaller one, then, given Scanlon's individualist restriction, any person in the smaller group could argue that this did not give her a fair chance of being saved. This person could argue that everyone deserves the same chance of avoiding harm and that a more appropriate principle would be one which required the agent to flip a coin in deciding which group to save. This would give each person in the situation the same 50% chance of being saved. Parfit calls this the equal chance principle. Neither Parfit nor Scanlon endorse this principle. They acknowledge that since the larger group could have many more people than the smaller group, it would be wrong for an agent to flip a coin and risk having to save the smaller group, letting many more people suffer harm. Parfit and Scanlon propose different solutions to this problem. Parfit believes that, as long as contractualism is allowed to consider the implications of a principle only for any single person, no person will be able to reject the equal chance principle. This is because cases involving groups of different sizes will be, according to contractualism, relevantly like cases where there is only one person in each group. Since, in cases of the latter sort, giving both people an equal chance of being saved seems appropriate, this will also be an appropriate way to handle cases of the former sort. The only way to avoid this, according to Parfit, is if the people in the larger group are allowed to argue that all of their reasons together outweigh those of the people in the smaller group. This would involve dropping the individualist restriction.⁷ It would allow different individuals' claims to be aggregated.

Scanlon has a different response. He believes that a person in the smaller group could not demand a fair chance of being saved because this consideration would have no bearing on the substantive burdens faced by the people in each group. According to Scanlon, whenever there are clear substantive burdens faced by individuals, these are the primary considerations which go into deciding who should be saved. Non-substantive burdens, such as risks of being harmed, are secondary to substantive burdens and should only be considered if the substantive grounds for saving one or the other group are unclear, for example if the harm that each group faced was not certain but only probable. The priority of substantive burdens is determined contractually. If a person will certainly suffer harm if she is not saved, this gives her a stronger personal claim to be saved than someone who faces, say, a 50% chance of suffering that same harm (all other things being equal).⁸ However, in the case of choosing who to save when the groups are uneven, being given a fair chance is, according to Scanlon, an irrelevant consideration since in either case some people will suffer substantive harm, and others will be spared from it.

Parfit could respond that even if the substantive burdens for each person in the situation were the same, this would not change the fact that none of these people could reasonably reject the equal chance principle. Since the ultimate stakes would be the same no matter which of the two principles (tie-breaker or equal chance) were followed, contractualism dictates that we must follow whatever principle is fairest for each person. The equal chance principle treats each person fairly. Thus, a person in the smaller group could reject the tie-breaker principle because it allowed her to be treated unfairly. Moreover, if, as a result of following the equal chance principle, the smaller group was saved and the larger group forced to suffer harm (perhaps as the result of an unlucky coin flip), none of the people in the larger group could reject the equal chance principle based on the fact that it did not take into account the value of saving her from harm. On the contrary, the equal chance principle respects these people by ensuring that each of them is given the same chance of being saved as everyone else.

Who has the better argument, Parfit or Scanlon, depends on a number of considerations. For instance, one is whether in giving everyone an equal chance of being saved, the equal chance principle actually disrespects the people in the larger group by not taking into account the full value of saving them from harm. If we thought that the equal chance principle did disrespect these people, we might ask whether being disrespected

in this way is worse than not being treated fairly in virtue of not being given an equal chance of being saved.

Therefore, it is unclear whether Scanlon's treatment of substantive vs. non-substantive burdens allows him to reject the equal chance principle outright. However, for the purposes of this paper, I will be more concerned with what the general implications are of Scanlon's hierarchy of burdens (substantive over non-substantive). I will be arguing later that always allowing a substantive burden to trump a non-substantive one has unacceptable consequences, and in addition, this hierarchy makes it unnecessarily difficult for contractualism to deal with cases involving numbers.

3. Scanlon's Argument and Its Ability to Deal with Different Harms

A much more glaring problem for Scanlon's tie-breaker view is its apparent inability to deal with cases where the harms faced by two different groups of people are unequal. According to the tie-breaker principle, numbers can only break ties in cases where each of the individuals involved faces the same burden. But what about cases where a much larger group faces a slightly lesser, but seemingly comparable harm to that faced by the people in the smaller group? There could be cases in which hundreds of people are in danger of being paralyzed, and on the other side, a case in which only one person faces premature death. It seems that, if the numbers of people facing paralysis were very great, we should save these people, even if it means letting one person die. Scanlon's tie breaker view does not allow this. It suggests that we should always save whoever faces the greatest burden. Is this even plausible? Scanlon does not believe so. Although he does not have a well worked out solution to these types of cases, he gives some tentative suggestions about how contractualism might deal with them:

If one harm, though not as serious as another, is nonetheless serious enough to be morally "relevant" to it, then it is appropriate, in deciding whether to prevent more serious harms at the cost of not being able to prevent a great number of less serious ones, to take into account the number of harms involved on each side. But if one harm is not only less serious than, but not even "relevant to," some greater one, then we do not need to take the number of people who would suffer these two harms into account in deciding which to prevent, but should always prevent the more serious harm.⁹

This seems like an ad hoc solution to the aggregation problem. Why

should the individualist restriction, which Scanlon sees as the guiding feature of contractualism, apply in some cases but not others? Whatever its merits may be, we will see later that this "relevance" approach is problematic. In any case, the challenge still remains. One thing a moral theory must be able to do us provide clear and decisive solutions to these problems. In my view, Scanlon has not yet shown that his contractualism can do this.

This brings us to contractualism's current situation. To recap, Scanlon has given an argument which does not allow the small benefits reaped by many to outweigh the serious harms suffered by a few. This is a result of the structure of contractualism, according to which the strongest individual claim trumps all other claims. He has also endorsed what he calls the tie-breaker principle, which asserts that when different groups of people face the same harm, we ought to save the larger group, since, if we did not, any person in that group could complain that her presence made no difference in how the agent went about deciding who to save. Finally, Scanlon has suggested that when certain harms are relevant to more serious harms, contractualism may allow agents to prevent the lesser harm, provided that the numbers of people suffering this harm are much greater than those suffering the slightly more serious harm.

4. Can the Aggregation Problem be Solved in a Non-Circular Way?

Many philosophers have argued that contractualism as a moral theory is circular.¹⁰ Joseph Raz in particular has tried to show that contractualism can only solve the aggregation problem in a circular way. Raz tries to show this by focusing on Scanlon's tentative suggestion that contractualism might be able to come up with a principle which would let an agent save many people from a severe harm, while ignoring one person who faced a slightly worse harm. Scanlon suggests that if an agent followed such a principle, a person in the larger group could reject this principle because

It did not give proper consideration to his admittedly less serious, but still morally relevant, loss. One might then argue that such an individual's claim to have his or her harm taken into account can be met only by a principle that is sensitive to the number of people involved on each side. I am not certain how such an argument would go, but it does not seem to me to be excluded in advance by the general idea of contractualism.¹¹

Here is Raz's response:

The problem is that Scanlon's contractualism does not exclude arguments of that form. It means that a person can object to a principle if, by disallowing aggregation where it is required, the principle does not give that person's harm the weight or role that it merits. And that means that the problem of aggregation has to be solved first, and contractualist arguments far from contributing to its solution cannot be deployed until it is solved.¹²

Here Raz is assuming that a principle still under consideration by the contractualist formula can disallow certain kinds of solutions that are in fact required, but this is not so. Contractualism does not assume anything about what kind of moral reasoning is required until the outcome of the contractualist test. This means that any principle under consideration could not "disallow aggregation where it is required." Any principle disallowing what is required would already have been rejected in reaching a decision about what is required. Raz could argue that, apart from any contractualist considerations, aggregative moral reasoning is required in some cases, but what is important to notice is that contractualism does not make such assumptions, and so is not circular in the way that Raz claims. This is true, not only of the example of aggregation that Raz discusses, but of contractualism's dealings with moral dilemmas in general. By considering only agent-relative reasons for rejecting principles, contractualism excludes an entire class of reasons from "affecting the outcome of the contractualist test."¹³ This is what Scanlon means by excluding impersonal reasons from the contractualist formula. Raz believes that, by excluding impersonal reasons, Scanlon's only aim is to show that claims about right and wrong concern people, and not the intrinsic value of "mountains, or the American condor."¹⁴ But this is not the whole story. It is true that contractualism does not consider these things, but that is because they do not involve reasons that are agent-relative. Non-agent-relative reasons concerning, not just mountains and condors, but people as well, are banned from consideration in contractualism. For example, according to contractualism, one could not reject a principle specifically because it made people's lives worse or because it brought about the worst outcome. These reasons do not reflect personal burdens. In order to be considered by the contractualist formula, reasons must make "an ineliminable (and non-trivial) pronominal back-reference to the person to whom the reason applies."¹⁵ This is what it means for reasons to be agent-relative. They must be in this format: "my burden", "my pain", etc.¹⁶

5. Is Contractualism Better Off Without an Individualist Restriction?

Although I believe Raz's argument rests ultimately on a mistaken conception of the structure of contractualism, he does pose a challenge to contractualism indirectly. The challenge is that if contractualism can solve moral dilemmas involving numbers, it must do so within the confines of the individualist restriction. Otherwise, it invites the circularity charge by allowing other considerations such as states of affairs, to be the wrong-making features of acts, while justification to others is simply an offshoot of this. This gives us reason to reject Parfit's suggestion mentioned earlier, that Scanlon should drop his individualist restriction in order to reach sensible solutions to cases involving numbers. This, however puts contractualism right back where it started in being unable to offer a sensible solution to cases in which a much larger group of people face only a slightly less serious harm than one or a few people. So it seems that Scanlon is faced with a choice between two evils. If he drops the individualist restriction this will make his theory vulnerable to circularity charges, but if he doesn't, he must bite the bullet and admit that saving one person from a more serious harm is always right regardless of how many others face harms that are only slightly less bad.

Biting the bullet is not an attractive option. As one writer puts it: "Can anyone who really considers the matter seriously honestly claim to believe that it is worse that one person die than that the entire sentient population of the universe be severely mutilated? Clearly not."¹⁷ This seems to me, an accurate reflection of most people's sentiments. Not being able to explain why we should save the entire sentient population from mutilation over one person from death leads us to ask why we should take Scanlon's theory seriously at all as a theory of what we owe to each other. On the other hand, dropping the individualist restriction would give us a solution to these cases, but would invite the charge that contractualism is circular or "vacuous" in Raz's words because it does nothing to establish the rightness or wrongness of saving or not saving one of the two groups.

Scanlon acknowledges this problem. He believes that whether or not contractualism would be a stronger theory without an individualist restriction depends on "the way in which elements of the resulting view restrict aggregative arguments."¹⁸ Here, Scanlon is talking about the restrictions on giving small benefits to a great number of people by imposing serious harm on a single person. He believes that in such

cases, we should always save the one person, no matter how many people stand to receive the small benefits. He considers the possibility of replacing the individualist restriction with a principle which blocked aggregation only in cases where the small harms or benefits are not relevant to the severe harm faced by the single person.¹⁹ This is the rationale behind what Parfit calls the triviality principle.²⁰ Scanlon does not try to develop this line of thought, but does not reject it either. Will replacing the individualist restriction with some kind of triviality principle solve the aggregation problem?

6. Problems with Establishing Thresholds of Relevance between Harms

Alistair Norcross has an argument which undermines the idea of a triviality principle. He claims that making some harms irrelevant to others would violate two very obvious and important doctrines involved in comparing harms.²¹ One is transitivity. This is the assumption that, given any three harms, (a), (b) and (c), (a) being the worst and (c) being the most trivial, the relevance of (b) to (a), and of (c) to (b), entails the relevance of (c) to (a). This is based on the assumption that there is

a descending scale of finitely many different harms, from the most serious, such as death, all the way down to the most trivial, such as a minor temporary headache. The difference is seriousness between any two adjacent harms is no larger than is necessary for the lesser harm to be clearly less serious than the greater harm...also...for every harm on the scale harm on the scale above the most trivial, there is some lesser harm that is relevant to it.²²

Norcross refers to this as the “continuity assumption.” He claims that transitivity and continuity together entail that the most trivial harm is relevant to the most severe harm. Moreover, it seems that there is no principled way to deny either one of these assumptions. Preserving transitivity by rejecting continuity would involve establishing a threshold of relevance between two harms on Norcross’s descending scale. In other words, we would have to say that the harm directly below the threshold is not relevant to the harm directly above it. But this seems impossible given that the difference between each of the finite harms on the scale is just small enough to make one clearly morally relevant to the other. Positing a threshold of relevance would be completely arbitrary.

Accepting continuity while denying transitivity would be equally absurd. This would entail that, in a case where we could prevent

only one of three harms, where the relevance between these harms is intransitive, we would be acting immorally no matter which harm we prevented. Norcross illustrates this with an example:

Suppose, for the sake of argument, (a) that the loss of both arms is less serious than but morally relevant to death; (b) that a broken leg is less serious than but morally relevant to the loss of both arms, but not morally relevant to death; (c) that in a choice between saving one life and preventing one thousand people from losing both arms, it is obligatory to aid the larger group; (d) that in a choice between preventing one thousand people from losing both arms and preventing one million people from breaking a leg, it is obligatory to aid the larger group. (The choice of examples is unimportant.) Consider now three different choices: (i) Save one life or prevent one thousand people from losing both arms. (ii) Prevent one thousand people from losing both arms or prevent one million people from breaking a leg. (iii) Save one life or prevent one million people from breaking a leg. From (b), (c) and (d) it follows that it is obligatory to aid the larger group in (i) and (ii) and the smaller group in (iii). So far, so good. But what happens when we are faced with all three options in one choice? No answer here seems satisfactory.²³

This is because, for each group that we could save, intransitivity suggests that we are morally obligated to save someone else. If we save the one person from death, we are doing something wrong since (c) tells us we are morally required to save the thousand from losing their arms. However, saving the thousand would be wrong since (d) tells us we are morally required to save the million from breaking their legs. Finally, even if we saved the million, we would be doing something wrong since (d) also tells us that broken legs are irrelevant to death and that we should, instead of preventing this harm, save the one person. This might be seen as a *reductio ad absurdum* of Scanlon's (admittedly tentative) "Relevance" argument. Moreover, if Scanlon dropped his individualist restriction, he could not prevent someone from using Norcross's argument as a reason for rejecting the triviality principle. Say for example that, in a choice between preventing the death of one person or giving extremely small benefits to millions of people, you decide to save the one person. Suppose that the principle you are following in saving the one person is the triviality principle. According to this principle, the small benefits that these millions stood to receive were simply irrelevant to the harm of the one person. Every person in this group of millions could then reject the triviality principle because it arbitrarily posits a threshold of relevance between harms, and therefore, unduly denies them benefits.

The point is that, if, in deciding whether or not the individualist restriction should be dropped, Scanlon's primary concern is how this will affect contractualism's ability to restrict aggregation of the kind just mentioned, he is better off keeping the individualist restriction. Without it, any principle trying to restrict the kind of aggregation Scanlon is against will involve positing arbitrary thresholds of relevance between harms.

And so, contractualism is back in a familiar place - not being able to solve cases involving intermediate harms. So far it seems that the structural changes necessary for solving that problem give rise to other problems that are equally challenging. Contractualism is not out of trouble yet. There is still a Scylla and a Charybdis between which it must navigate. I turn now to the issue of how this can be done.

Part II: Contractualism's Answer

7. The Role of Generic Reasons in Contractualism

An important feature of contractualism is that, in order to determine whether a principle could be reasonably rejected by others, the standpoints which we must consider include not only those to whom justification is owed, that is, those who are affected most directly by our actions, but anyone who might or might not be able to reasonably reject the principle.²⁴ We must consider, not only the "consequences of particular actions, but also the consequences of the general performance or non-performance of such actions and the other implications (for both agents and others) of having agents be licensed and directed to think in the way that that principle requires."²⁵ Why is this so? One of the most obvious reasons is that the general authorization (or prohibition) of certain actions have implications which extend beyond the actual consequences of the actions of agents when they are performed. Because of this, we have to consider what consequences a principle would have if it were generally accepted. Scanlon claims that in doing so, "our assessment cannot be based on the particular aims, preferences and other characteristics of specific individuals."²⁶ This is so because "we cannot know, when making this assessment, which particular individuals will be affected by it in which ways (who will be affected by it as an agent required to act a certain way, who as a potential victim, who as a bystander, and so on)."²⁷ Instead, we must rely on commonly available information about what people have reason to want. Reasons of this sort are what Scanlon calls generic reasons.²⁸ In this section I will argue that, in addition to being an important part of the structure of contractualism,

generic reasons are an important part of solutions to problems involving aggregation.

One might ask, if generic reasons are such an important part of contractualism, especially in helping it resolve a glaring problem, why have they not been brought up until now. The answer has to do with Scanlon's view about the kinds of cases involving numbers we have discussed so far. Scanlon believes that if the occurrence of a particular type of case is sufficiently rare, we do not, in considering what principles to follow in those cases, have to give much consideration to what implications those principles would have if generally followed.²⁹ Since the kinds of examples discussed in the first half of this paper are hypothetical and are fairly uncommon, Scanlon does not consider the aspect of generic reasons which, I will argue, is crucial for solving the aggregation problem.

The rareness of particular cases should not influence our consideration of generic reasons. Contractualism is a hypothetical contract theory about morality. It asks what principles anyone would accept, given the aim of finding such principles. As such, any example it considers should be viewed as hypothetical, whether its actual occurrence is common or not. How frequently certain kinds of cases occur may enter as a consideration of what rules should govern certain societies but this rationale seems rather conventionalist and rule-consequentialist in nature. It should not determine how any particular moral dilemma should be solved within a contractualist framework. This is the first substantial difference between Scanlon's view and my own. My contention is that, in discussing particular cases involving numbers, we must consider what implications any principle we use to solve those cases would have if generally followed.

8. Generic Reasons and Numbers

Generic reasons, I believe, add a new dimension to our understanding of cases involving numbers which is generally overlooked. When considering such cases, most people consider only the consequences for the people involved in that particular case. This is an understandable reaction given that, in cases where different groups of people are facing some harm, the numbers are the most salient feature of the case, and it seems that appealing to the numbers is good enough to determine the rightness of saving the larger group. What a contractualist must argue is that such problems are not mere math problems. Numbers do

not provide a bare account of the rightness of saving the larger group. A principle which relies on numbers will only be acceptable if no individual has a personal reason for rejecting it. Once it is acknowledged that justification to others counts twice; that is, in terms of the burdens imposed on the people who are most directly affected by the actions of an agent, and in terms of what individuals in general would have reason to want, a solution to problems involving aggregation presents itself.

The model I have in mind is roughly this: When confronted with situations involving aggregation, first, we have to consider the possible consequences of accepting or rejecting some principle for each person in the specific example. In doing so, we must hold true to the individualist restriction. We must consider only the reasons that individuals have for accepting or rejecting some principle. Second, we must consider the consequences for everyone of such a principle if it were generally followed. If such a principle would, in general, make things worse for individuals, then this is one consideration in favor of rejecting the principle. Again, we must be careful not to wander outside the confines of the individualist restriction. These two considerations make up the generic reasons in favor of or against certain principles. Finally, we must weigh all the reasons people have for rejecting a principle to see whose claims are actually the strongest. We must decide whether the generic reasons that an individual has for rejecting the principle are stronger than the generic reasons that any other individual has for accepting it.

This brings us to the second substantial difference between my view and Scanlon's. In order for Scanlon to accept this model, he will have to give up the view, discussed at the outset of the paper, that substantive burdens always trump non-substantive ones. This is because, as we will see, when we consider what implications a principle would have if it were generally followed, one particularly important consideration will be how it affects a person's chances of being saved if he ends up in a certain situation. Revising this part of Scanlon's view in order to solve his aggregation problem is not an ad hoc maneuver. There are already good reasons to deny this precedence of substantive over non-substantive burdens. One of them is that there could be cases where, on one side, a single person faces premature death and on the other, millions of people each face a 99% chance of dying prematurely. According to Scanlon's hierarchy of burdens, we should save the one person, since she is the only one who faces a substantive burden. The others face the non-substantive mere chance of a burden. But this cannot be right. If we saved the one person in this case, it is almost certain that millions of people would die. A theory which allowed this would be as

implausible as a theory which allowed everyone in the universe to be mutilated in order to save one person. If Scanlon allowed substantive and non-substantive burdens to be considered concurrently, instead of one having precedence over the other, he could avoid this implausible consequence, and also, as I will show, offer a better overall answer to his aggregation problem.³⁰

9. Cases

In order to see how the abovementioned method can be used to solve the aggregation problem, we need to see how it applies in specific examples. Here I will be considering both examples where the numbers of people in each group are unequal but the harms they face are the same and cases where the harms faced by each group are unequal.

First, consider:

Shipwreck. A ship explodes, sending three people flying in different directions. Two people land very close to each other, but the other person lands far away from these other two. As the captain of the rescue boat, you have two options. On your first option you can throw your (only) life preserver to the two people. If you do, these two people will be saved, but the other person will drown. On your second option you can throw the life preserver to this person, but if you do that, then the other two people will drown. What should you do?

Here it is fairly obvious that you should save the two people. What (may not) be so obvious, is what generic reasons people have for accepting a principle allowing that action. To determine what those reasons are, we should consider what, if agents were generally licensed to save the larger group, the chances would be that anyone who might end up in a situation like Shipwreck would have of being saved. In order to do this, consider any three (randomly selected) people who might end up in a similar situation. Since anyone could end up in this kind of situation, we should assume that 'people who might end up in a similar situation' picks out everyone.³¹ Since the chances of any given person out of three ending up in the smaller group is $1/3$ and for that same person, the chances of ending up in the larger group is $2/3$, each person would have a $2/3$ chance of surviving if agents were generally required to save the larger group in such cases, but only a $1/3$ chance of being saved if agents were required to save the smaller group. Since people in general would have reason to want a principle which gives them a better chance of being saved, such a principle, it seems, could not be reasonably rejected

by anyone. One should notice that the single person in the smaller group also has these reasons, and although the harm he faces strengthens his claim to be saved, if you were to save him and let the two drown, he could argue that, in saving him, you acted on a principle which actually gave him a worse chance of being saved, given his situation. It was only by luck that he ended up in the smaller group and was saved.

This example shows how contractualists could determine what agents who are faced with such situations are morally required to do. Those moral requirements will, of course, change depending on the specific features of each particular case. There may be cases in which the person who ends up in the smaller group had a greater chance of ending up by himself. In cases like Shipwreck, this is not true. In that case there is nothing different about any particular individual that would make it more likely that she would end up in the smaller group. But it does seem possible that this might happen and, when it does, whether the person in the smaller group should be saved instead of the people in the larger group will depend on whether everybody in general would have reason to accept a principle which allowed this. To see how these considerations might yield different results, we need to consider some more examples. Suppose that,

Sharks. Three surfers fall off their boards and will be eaten by sharks unless rescued. However, because one of these surfers is far away from the other two, you, the captain of the rescue boat, only have time to save one group. However, there is an additional feature to the case: because this one surfer frequently preferred to surf alone, instead of with friends, his chances of being alone in a single group were greater than his chances of being in the larger group. Who should you save?

Again, here you should save the larger group. The fact that the surfer frequently surfed alone does not seem to affect this decision in the slightest. Do the generic reasons favor this answer? One might argue that the generic reasons which favored your saving the larger group in Shipwreck cannot be the reasons in favor of saving the larger group in Sharks. In this case the single person in the smaller group could not argue, if you were to save him, that the principle allowing this gave him a worse chance of being saved. This is because his chances of ending up in either group were not the same as everyone else's. His preferences gave him a better chance of ending up alone, and therefore he has a stronger claim than anyone else to be saved. This line of argument, however, is mistaken. This is because, when we consider what implications a principle would have for everyone if generally followed, lone surfing is not part of the body of commonly available information

about what people have reason to want. This body of information includes things like certain basic rights, health care, etc. These are things that anyone would have reason to want.³² However, not everyone would want special privileges for lone surfers. This consideration is too particular to count as generic. So a principle allowing you to save him based on his preference for lone surfing is not what people in general would want. People in general would not want a principle which required an agent to save the lone surfer unless lone surfing was a trait commonly shared by all people, or perhaps had some commonly understood importance to people. But, since this is not the case, the surfer cannot use his situation as a reason for our saving him. Instead, the generic reasons favor the larger group. In general, a principle requiring the agent to save the larger group would protect everyone better, since if they were to end up in that situation, they would have a greater chance of being saved. However, there are cases where the generic reasons do favor the person in the smaller group. For example, suppose that:

Organs. Five patients will die unless they receive a much needed organ. One needs a heart, another, a kidney etc. You, the doctor have a choice to drug one of your healthy patients who is in the hospital for a routine checkup (and who happens to have the right tissue compatibility) and distribute his organs to each of these five patients. If you do, you will take one life, but save five. Otherwise, the one healthy patient will live, but the five ailing patients will die. What should you do?³³

This example has moral implications which are very different from those in either Shipwreck or Sharks. First of all, since people in general are more likely to be healthy patients than patients who need an organ transplant, and because things such as routine checkups are part of almost every person's life, any person would, in general, have reason to want a principle which prohibited an agent from killing the healthy patient. Perhaps this principle could be stated in terms of protecting the vulnerable. There is also a very general demand that doctors in particular not be licensed to harm the patients whom they are supposed to help. So the generic reasons against you killing the one to save the five include both the fact that people in general would not want agents to behave in this way, and a reason against having doctors in particular be licensed to do that kind of thing.³⁴ These generic reasons seem strong enough to create a moral requirement that the doctor not kill the healthy patient.

So far, I have suggested how contractualism can solve certain cases where different groups of people faced the same harm. An attractive

quality of this method is that it can also be used in cases where the harms faced by the people in each group are different. Consider for example, a case in which you could either save one person from certain death or two people from paralysis.³⁵ If you saved the two people from paralysis, the person in the smaller group could object that this imposed a certain burden on him, the burden of premature death. But anyone could object that if a principle allowing the agent to save the two from paralysis were generally implemented, this would, in general, give everyone a worse chance of a lesser burden. This is because, any one person out of three who might end up in this situation³⁶ would have a $1/3$ chance of ending up in the smaller group by himself and face premature death, but a $2/3$ chance of ending up in the larger group facing paralysis. Assuming that a $2/3$ chance of paralysis is better than a $1/3$ chance of death, any of these people would want a principle giving them this chance. Since such a principle requires the agent to save the one person, this is what he must do. Also, since this one person faces a greater harm than either person in the larger group, he has a stronger generic reason for wanting to be saved than either of these people. Since the generic reasons favor saving the one person, this is what is morally required of the agent.

However, what the agent is morally required to do changes if the numbers of people suffering the slightly lesser harm are great enough. If, for example, the choice is between saving one person from death or one hundred people from paralysis, the agent should save the larger group, for if agents generally acted in this way, any one person out of any hundred who might end up in such a situation would suffer the comparatively less bad $1/100$ chance of death instead of a $99/100$ chance of paralysis. In this case, although the substantive harm of the person in the smaller group is greater than anyone else's harm, it is not enough to outweigh both the substantive harms of any person in the larger group and the abovementioned generic reasons in favor of saving these people.³⁷

A more challenging type of case for contractualists arises when the harm that a single person faces is much worse than the harms which a much larger group of people face. Consider, for example, a choice between letting an innocent person, let's call him Jones, die so that a million people can be spared from some small harm such as a slight headache for an hour,³⁸ or saving Jones and allowing the million to suffer the headaches. Many will feel that it would be horrible to let Jones die so that we could spare any number of other people from a harm which, in comparison, seems insignificant to the one Jones faces. Do the generic reasons favor our saving Jones? One might argue that they are actually

in favor of letting him die since, a principle generally allowing this would mean that for any one person out of any one million people who might end up in this situation, this person would have only a $1/1,000,000$ chance of dying, but an almost certain (999,999 to 1) chance of suffering a slight headache. Which one is more preferable? Norcross suggests that it would be worse to suffer the slight headache. He invites us to imagine the following case which I will call:

Headache. You are settling down to spend the next twenty-four hours at home, reading, watching movies, eating and sleeping, when you feel the onset of a moderate headache. You know from experience that this headache will last for twenty-four hours, unless you take your favourite brand of pain-killer. Alas, the medicine cupboard is bare. However, the nearest pharmacy that sells your brand is only three miles away, less than a ten minute trip in your car. So, you jump in the car, purchase pain-killer, and spend a pain-free twenty-four hours. Were you irrational to do that?...Of course not. Suppose we add the following detail. You have just read an article in a reliable publication that claims that the type of car journey you are considering increases your chances of death over staying home, by one in a million. Does that change our original judgment about the rationality of your action? No. Most, if not all, of us were already aware of the risks of traveling by car, when we made our original judgment. Many of the things we do to improve our quality of life involve similar small risks of death [and]...many of them, such as driving (or walking) to the cinema to see a good movie...are clearly rational.³⁹

Since most of us routinely take such risks, they do not mean that much to us. Most of us, however, would want to avoid pain even if that pain were very slight. The generic reasons then, seem to favor letting Jones die since, any principle which allowed that would give everyone a better chance of a more preferable outcome. Are we morally required, then to let Jones die? We are not, because when we weigh the generic reasons of each individual in both groups as well as the generic reasons applying to everyone in general, Jones' burden (death) gives him a generic reason which outweighs both the generic reasons of any person in the larger group (a slight headache for an hour) combined with the generic reason (an almost certain chance of having a slight headache for an hour) which any individual who might end up in that situation would have if a principle allowing us to save Jones were generally implemented. Since the considerations in favor of saving Jones outweigh all other considerations, we are morally required to save Jones.

This method solves particular cases of the kind just mentioned in an intuitively appealing way. It is consistent with the intuition that any

number of slight headaches would not be enough to create a moral requirement for an agent to let the one person die in order to prevent these headaches. We can see that, as the number of people with headaches increases, the substantive burdens of any one person in that group remains the same, and the generic reasons which anyone could offer against our saving Jones do not change in any way that would be significant for any individual. For example, if the number of people facing slight headaches were increased to a billion, a principle which generally allowed the agent to save Jones from death and let these billion suffer their headaches, would mean that any person, if they ended up in that situation, would have a 999,999,999 to 1 chance of suffering a slight headache. Since the difference between this chance and the 999,999 to 1 chance (which we saw in the previous example), of suffering a slight headache makes almost no difference to any particular individual, the increase in numbers does not change the moral requirement to save Jones. In fact, as the number of people suffering slight headaches approaches infinity, the chance which any person, who might end up in such a situation, would have of suffering a slight headache approaches 100%. But this does not make any difference since, even if it could reach 100%, this would still not be a strong enough reason to ignore Jones, since a 100% chance of suffering a slight headache for an hour does not give anyone a reason that can outweigh Jones' reason.

We also see that there is nothing about this solution which violates the transitivity involved in comparing different kinds of harms. We saw in section 6 that a principle which allowed an agent to save Jones which relied on the claim that small harms like slight headaches, are not relevant to greater harms such as death, denies the transitivity and continuity assumptions, and would involve an arbitrary threshold of relevance between harms. But the argument I have given does not violate transitivity or continuity. A contractualist can even admit that, if the numbers of people suffering slight harms were great enough, letting an innocent person die in order to spare them from those harms might be better than allowing them to suffer the slight headaches. But since contractualists do not rely on the goodness of outcomes to solve moral problems, this fact would not, according to contractualism, be enough to determine what we ought to do in those situations.

Here I will address one more worry. Scanlon considers the idea that if cases like the one involving Jones arose frequently, the masses of people suffering headaches could appeal to what he calls intrapersonal aggregation — the aggregation of small inconveniences within an individual life, in order to outweigh the reasons a single person facing

death might offer for our saving him.⁴⁰ We can certainly imagine possible worlds where we are frequently faced with the choice to save one person from death or billions from slight headaches and if we kept allowing the headaches, billions of people would continually suffer throughout their lives. Scanlon is worried that this kind of situation would fall within the range of cases where the aggregated pains of the many are relevant to the severe burdens of the few, and since we have to consider, in each individual case, what the consequences of our actions would be if they were regularly performed, one might think that the right thing to do in a case like that of Jones would be to let him die and prevent the headaches. Of course Scanlon rejects this conclusion since he does not think cases like Jones arise frequently enough for us to consider what would happen if a principle favoring Jones over the masses were frequently implemented. However I do not have this luxury. I have claimed that the actual frequency of hypothetical cases should not influence how we consider which principles are acceptable and so it may seem that I am committed to the implausible view that we should always save the single person from death even if this will mean constant suffering for the masses.

However, I don't think I am committed to this. To see why, first, imagine one scenario in which such cases arise frequently, but where the billions of people suffering slight, one-hour headaches are different each time a dilemma arises. Suppose, for example, that after saving Jones and letting everyone else in the world suffer slight headaches with no other negative consequences, another person, Smith is in the same danger. However, this time, saving Smith will not mean more pain for everyone else here on Earth. Instead, let's suppose that saving Smith will mean slight one-hour headaches for billions of people on planet X. As long as the headaches that result are suffered by different people, it seems plausible that we should save the single person from death. Thus we can frequently follow the principle that would allow us to save Jones without worrying about the intrapersonal aggregation of the masses giving them a reason to reject that principle. Scanlon's worry is the result of his confusing a principle which, if followed regularly, would have accruing negative consequences, and a principle which explicitly allows those negative consequences. The latter would be a principle which allowed the same people to suffer continuously by adding small harms to their already diminished welfare, and it seems like anyone would have good reason to reject such a principle.

10. Conclusion

I have argued that contractualism can deal with cases involving numbers if it considers the general implications of principles for everyone, not just in cases that are common, but in any hypothetical case. It also must allow substantive and non-substantive burdens to be considered together, instead of giving substantive burdens precedence over non-substantive ones. The method I have suggested for solving Scanlon's aggregation problem is based on the idea of acting in accordance with principles which any individual would have good reason to want. Since, in general, a reasonable individual would want to have a better overall chance of being spared harm, it would be reasonable for her to reject a principle that would reduce her chances of being saved from that harm. An example of such a principle would be one which allowed a potential saver to save a smaller group of people rather than a larger group from suffering the same harm. Since this principle would, if generally followed, reduce the chances of being saved for everyone, every individual has reason to reject it, and accept an alternative principle which increases their chances of being saved. This way of dealing with numbers also covers cases where the harms faced by different groups are unequal. Here, the consideration that any individual would want a better chance of suffering a lesser burden also applies. The general acceptance of a principle meeting these standards requires that a potential saver act in different ways depending on the specifics of his situation. He may be required to save the people in the larger group if the harm they face is not as bad as the harm that those in the smaller group face, but still quite serious. He may be required, on another occasion, to save the smaller group if their harms would be much more serious, or if the difference in numbers between the groups is not very great. This method, rather than relying on states of affairs or on impersonal reasons, relies on principles that no individual would reasonably reject. This makes it a solution that is intuitively appealing and, at the same time, consistent with the basic idea of contractualism.

References

- Gauthier, David. (2003): "Are We Moral Debtors?," *Philosophy and Phenomenological Research* 66 pp. 162-168.
- Gibbard, Allan (2003) "Reasons to Reject Allowing," *Philosophy and Phenomenological Research* 66: pp. 169-175.

- Harman, Gilbert (1997): "Ethics and Observation," in *Moral Discourse and Practice: Some Philosophical Approaches*, ed. Stephen Darwall, Alan Gibbard, Peter Railton (New York: OUP).
- Kamm, F.M. (2002): "Owing, Justifying and Rejecting," *Mind* 111, pp. 323-354.
- Miller, Richard W. (2002): "Moral Contractualism and Moral Sensitivity: Critique and Reconstrual." *Social Theory and Practice* 28:2 pp. 193-220.
- Norcross, Alistair (1999): "Great harms from small benefits grow: how death can be outweighed by headaches." *Analysis* April 1998: pp. 152-158.
- Norcross, Alistair (2002): "Contractualism and Aggregation" *Social Theory and Practice* 28:2 p. 303-314.
- Parfit, Derek (2004): "Justifiability to Each Person" in *On What We Owe to Each Other*, ed. Philip Stratton-Lake. (Malden: Blackwell Publishing) pp. 67-89.
- Pettit, Philip (2000): "A Consequentialist Perspective on Contractualism" *Theoria* 66 pp. 228-236
- Raz, Joseph (2004): "Numbers With and Without Contractualism" in *On What We Owe to Each Other*, ed. Philip Stratton-Lake. (Malden: Blackwell Publishing).
- Ridge, Michael (2001): "Saving Scanlon: Contractualism and Agent-Relativity" *The Journal of Political Philosophy* 9 pp. 472-481.
- Scanlon, T.M. (1998): *What We Owe to Each Other* (Cambridge: Harvard University Press).
- Scanlon, T.M. (2004): "Replies," in *On What We Owe to Each Other*, ed. Philip Stratton-Lake. (Malden: Blackwell Publishing).
- Suikkanen, Jussi (2005): "Contractualist Replies to the Redundancy Objections," *Theoria* 71 p. 38-58.
- Thomson, Judith Jarvis (1990): *Realm of Rights* (Cambridge: Harvard University Press) .

Notes

1. Scanlon (1998) p. 197
2. See Scanlon (1998) p. 142.
3. I will generally be using the word 'Contractualism' to mean Scanlon's Contractualism.
4. Ibid. p.230.
5. Norcross (2002), Miller (2002), Parfit (2004), Raz (2004).
6. See Parfit (2004).
7. F.M. Kamm has made this same suggestion for similar reasons. See her "Owing, Justifying and Rejecting" *Mind* 111: 323-354.
8. Scanlon elaborates on this idea in his reply to Parfit (2004).
9. Scanlon p. 239-240.
10. Gauthier (2003),Gibbard (2003), Kamm (2002), Pettit (2000), Raz (2004) Thomson (1990).
11. Ibid. p. 240-241.
12. Raz p.66 [*italics added*].
13. Raz, p.59.
14. Ibid.
15. Ridge, p. 475.
16. A point emphasized by Suikkanen (2005).
17. Norcross (2002) p. 307.
18. Scanlon (2004) pp. 131
19. Ibid.
20. Parfit, pp. 85.
21. Norcross (2002), (1999) .
22. Norcross (2002) p. 307.
23. Ibid. p. 308.
24. Scanlon (1998) p. 202.
25. Ibid.
26. Ibid.
27. Ibid.
28. Some confusion may arise here as to how agent-relative reasons based on a principle's particular effects on an individual differ in kind from reasons relying on concerns such as how people would be affected if that principle were generally accepted. This distinction, however, is misleading. Generic reasons include both: a principle's effects on each individual and the general consequences of this principle for persons in general. Whenever we consider possible reasons for rejecting or accepting principles, eligible reasons will include both of these considerations. And so generic reasons can be seen as a conjunct of the two.

29. Ibid p. 238
30. In discussing the way in which substantive and non-substantive burdens strengthen reasons for or against principles, I will be appealing to our intuitions. A more thorough account of how these different kinds of burdens interact is an issue worth pursuing but is beyond the scope of this paper.
31. One might object that people who never travel on boats do not count as part of this class, but that is not true. Even if a person has never traveled by boat, and does not plan on it, this does not mean that he could never end up in such a situation.
32. The serious skeptic will ask how we could possibly know what is on this list. Is there always a way of determining what it would be most reasonable to want? I believe that, ultimately, this list will have vague boundaries and there will be difficult boundary cases. However, hard cases seem to pose a challenge for every theory. There is no reason why they should pose a unique problem for contractualism.
33. I take this example from Harman's 'Ethics and Observation' in *Moral Discourse and Practice: Some Philosophical Approaches*, ed. Stephen Darwall, Alan Gibbard, Peter Railton (New York: OUP, 1997).
34. Why this distinction? Well, as it turns out, the fact that doctors in particular should not harm healthy patients is not the only generic reason one could offer for him not cutting up the healthy patient. Even if some random person who never took the Hippocratic oath, and had read a handbook on how to surgically remove organs sneaks into the room and kills the healthy patient, anyone could reject a principle allowing this since any person would, in general, have more of a chance of ending up in the healthy patient's situation than in the shoes of one of the five unhealthy ones. So I think there is a good reason to mention the distinction.
35. In this case, and from this point on, the reader should assume that the nature of the cases discussed is similar to that of Shipwreck or Sharks, in that there are no particular features, as there were in Organs.
36. Again, here I am assuming that this means everyone.
37. Again, here I am relying on the intuition that the fact that a principle allowing agents to save the single person from death, instead of 100 people from paralysis, would give anyone who ended up in that situation a 99% chance of being paralyzed, is enough to overcome the fact that the one person faces a harm which is slightly less severe.
38. An example I owe to Alistair Norcross. See his 'Great Harms from Small Benefits Grow' *Analysis* April, 1998: p. 152-158.
39. Ibid. p. 154.

40. Scanlon does not actually consider Norcross's headache case, but one similar to it involving an accident in a TV transmitter room. We are asked to imagine whether it would be right to stop the TV show for fifteen minutes in order to save the single person from an hour of severe pain, or allow the millions watching TV to continue enjoying themselves. See p. 235 of *What We Owe to Each Other*.

AN INTERVIEW WITH MICHAEL WALZER

Winter 2006
Princeton University

Each year The Dualist includes an interview with a modern philosopher chosen by the staff. This year, Michael Walzer graciously agreed to answer questions posed by The Dualist and by the Stanford Philosophy Department. Michael Walzer is editor of the political quarterly Dissent. He has become a significant advocate of just war theory, in opposition to members of the left who have neglected the theory. Walzer authored Spheres of Justice, which modified Rawls' theory of individualized justice, instead claiming that justice is a social concept, germane to the cultures and traditions in which it emerges. A universalized theory of justice does not satisfy Walzer: the specifics of justice vary from one culture and from one era to the next.

Debra Satz:

Whether and how do the arguments in Just and Unjust Wars apply to the "war" on terror?

Michael Walzer:

Insofar as the "war" on terror is mostly police work, they don't apply: the relevant issues have to do with civil liberty and executive authority, not with justice in the sense that the word has in just war theory. But there are useful comparisons to be made between the "rules of engagement" for soldiers and for police—for example, the rules with regard to innocent bystanders are stricter for police than for soldiers in battle, and it would be interesting to ask why that is so. But when the "war" on terror takes the form of a real war, as in Afghanistan, then all the arguments about just and unjust wars apply.

Satz:

On your theories in Spheres of Justice, what do you think of the recent Supreme Court decision on campaign finance?

Walzer:

I have a mixed view of campaign finance. On the one hand, it is important to make sure that money doesn't distort the results of

democratic elections (as it commonly does), and this clearly requires fairly strict regulation of campaign contributions. On the other hand, the giving of time, energy, and modest amounts of money to a candidate or a party is an important way of registering the intensity of one's commitment, and I would not want to shut that down. I don't think that the Supreme Court's decision fits either of these views.

Satz:

As a more general question, I'm interested in your thoughts on the justification of inequality. Where and when in your theory is inequality justified?

Walzer:

I am not bothered by any distributive inequalities within the different spheres of justice so long as these are determined by the appropriate distributive principles and so long as they don't distort distributions in other spheres. But obviously radical inequalities in one sphere, the market especially, do distort other distributions. And then we have to argue about how to deal with the distortions. In the case of the market, blocked exchanges are one way; redistribution through the tax or welfare systems is another. But if someone with a lot of money enjoys only the things that money can rightly buy, his wealth doesn't bother me. Should it? Should I be offended by it? It seems to me that we should be concerned with the uses of inequality, not with inequality itself.

Rob Reich:

The introduction to your best-known book Spheres of Justice says that the book emerged from a course you taught with Robert Nozick at Harvard around the time that Rawls's Theory of Justice was published. What happened at Harvard in the late sixties and early seventies that produced such important and long-lasting political philosophy?

Walzer:

Except for the fact that Rawls happened to be there, I don't think anything happened at Harvard that didn't happen everywhere else. It was the political ferment of those years that caught us all up—even people like Bob Nozick who weren't really "political people" (Bob was much more interested in deep philosophical questions that I could never get my mind around). The course that he and I taught was so exciting because of the intense interest of the students who came—and they were

interested not only in the fun of the argument but in the actual issues. It was their interest that drove the two of us to work very hard.

Reich:

You were the doctoral supervisor for Susan Moller Okin, the recently deceased political theorist from Stanford. She credited you with giving her the confidence to write a dissertation on the place of women in western political theory. How do you assess her legacy? Why were you enthusiastic about her project when so few other people were?

Walzer:

I had a number of excellent women students in those years— Amy Gutmann and Nancy Rosenblum as well as Susan. And I was as supportive of them as I could be because I recognized in them a certain kind of drive, ambition, commitment that wasn't going to be denied. They reminded me of an earlier generation of Jewish men. I wanted to be on their side.

Susan's *Women in Western Political Thought*, along with Jean Elshtain's *Public Man, Private Woman*, and Carol Pateman's *The Sexual Contract* transformed the way the history of political theory is written and the way theory itself is written. The impact was enormous, in large part because there were so many bright, ambitious, and often angry young women just waiting for the breakthrough those three made.

Reich:

You help publish Dissent; Josh Cohen helps publish Boston Review. In your mind, how does Dissent relate to your work as a political philosopher?

Walzer:

They are the same thing (which probably means that I am not much of a philosopher). I write an article about some political question that interests me, and after it's written I decide whether to publish it in *Dissent* or in an academic journal. If I decide for the latter, I have to add notes, references to the academic literature, and perhaps qualify the argument a bit, but I don't do anything more. I write in plain English; I don't like academic jargon; I have no head for abstraction. Writing in *Dissent* is more natural for me, but I can manage *Political Theory* or even *Philosophy and Public Affairs*.

Reich:

Woody Allen remarked that the editors of Dissent and Commentary should get together and publish a new magazine. It would be called Dysentery. Any comments on that joke?

Walzer:

It is a good joke. I don't think that it helped our sales much.

Ed Bruera:

The termination of civil wars has become an important issue for social scientists. The complex empirical and ethnographic issues surrounding this topic have been well examined, yet there has been surprisingly little normative work on the problems that tend to arise from this (for example, on the demobilization of former soldiers, especially child soldiers). What do you think the important ethical and moral considerations are in dealing with failed states and civil wars, as concerned observers and perhaps interveners?

Walzer:

Except with regard to humanitarian intervention, I have not written about these questions, and I am not sure that I have much to say about them. Failed states are obviously moral as well as political disasters, and so they teach us how important the state is: it's the only agency we have that, when it is competent and decent, can defend human rights. It follows that the international community should do everything in its power to avoid state failure. And it should do everything in its power, in the case of failure, to repair political institutions by whatever means are available—including, in extremity, military means. Warlords, child soldiers, rape and terror—you don't need a moral philosopher to explain the wrongness of all that.

Peter Stone:

Professor Walzer, you have argued that when the United Nations proves ineffective at stopping atrocities around the world – as it often does – the United States is justified in intervening unilaterally (presumably in violation of the UN Charter, which forbids unilateral military interventions as tantamount to aggression). But at the same time, few nations in the world are as hostile to the UN as the US. The US vetoes more than any other country, withholds dues, regularly browbeats UN officials, etc. Isn't there a serious moral problem raised when a country obstructs the UN, thereby rendering it ineffective, and then uses that ineffectiveness as a justification for doing what it wants?

Walzer:

No, I haven't argued that the US is justified in intervening unilaterally in cases of mass murder or ethnic cleansing; I have argued that *anyone* is justified in intervening in such cases. The examples that I usually refer to when I defend unilateralism are the Vietnamese intervention in Cambodia (to shut down the Khmer Rouge killing fields), the Indian intervention in East Pakistan (to end the terror there and to allow millions of refugees to return), and the Tanzanian intervention in Uganda (to overthrow the murderous regime of Idi Amin). The UN did not authorize intervention in any of those cases; nor would it have done so, had it been asked. Yet all of them seem justified to me. I think that the US, and other countries too, should be working to create an effective UN, which could deploy military forces capable of acting forcefully (as no UN force has yet done) in parts of the world where such forces are needed. But since a UN of that sort doesn't exist, someone else has to act, when forceful action is necessary. I don't see why the US should be excluded from acting, but I certainly prefer a division of labor, among as many countries as possible.

The Vietnamese invasion of Cambodia — my preferred example — certainly fits the UN's definition of aggression — so much the worse for the definition. No one in the world, who knew about the killing fields, would argue that Vietnam was a criminal aggressor. So we need to judge every intervention independently, support some, oppose others. And I don't see how a vote at the UN (even if the US were paying its dues and never using the veto) could possibly determine which ones we should support and which ones we should oppose. Will there ever be a UN whose decisions would be authoritative in that sense? No, we will always be on our own.

Robert Adcock:

How has the institutional and intellectual environment of the Institute of Advanced Study influenced the character and direction of your thought and scholarship since your arrival there in 1980?

Walzer:

My intellectual and political formation was pretty much complete when I came to the IAS. The only new direction that my work has taken since 1980 is toward Jewish studies. The big collaborative project on which I am currently engaged, on the Jewish Political Tradition, required a major educational effort on my part (with help from my friends), which I probably could not have undertaken if I were still teaching at a university. But that was largely a matter of time; it was not determined in any way by the intellectual environment of the Institute. I was already

very much engaged with Jewish politics before I came here.

Adcock:

How does the audience(s) you would ideally hope to engage compare to the readers you actually get? Has this changed over the course of your career?

Walzer:

No writer gets as many readers as he wants—at least no academic writer does. I have always imagined myself writing for a general audience, for the sort of people who read political magazines, the *New York Review of Books*, even monthlies like *The Atlantic*. And I have in fact written for that audience, but only intermittently. Mostly, I am read by college students who are assigned a chapter or two from one or another of my books—and then by some fraction of the liberal-left public (the social democratic fraction, maybe).

Adcock:

Would you welcome, reject, or qualify a description of yourself as a “public intellectual”?

Walzer:

It is an honorable name; I am flattered to be called a public intellectual, though I rarely reach a sufficiently sizable public to merit the name. Similarly, I am flattered to be called a social critic, though a real social critic would spend more time than I do criticizing our society (it needs it). There is an old-left image of the intellectual as someone with an all-encompassing theory, with which he/she can address any political issue that arises. I don't have a theory like that; my commentary on public issues, in *Dissent*, for example, is pretty ad hoc. And I probably spend as much time editing other people's comments as I do writing my own. I am a public editor (for a very small public).

From The Dualist Staff

Topic One:

One currently topical problem that relates to two strands of your work involves the intersection between cultural communities' values and just war theory. In traditional just war theory, as you have defined it, soldiers for both the just and unjust sides are equally subject to the requirements of

proportionality and decency that disallow the use of disproportionate force, the causing of unnecessary harm, and the torture and rape of enemy soldiers.

However, imagine a case where a force of American soldiers (in present-day Afghanistan, for example) meets a force of devoutly religious Afghan soldiers in battle. Imagine that the Afghan soldiers value their religious purity more than their own lives or bodily integrity – they would rather be raped or tortured than lose this purity. Also, imagine that this religious purity is dependent on some religious object, which is carried along with the Afghan battalion, to not be damaged by the American soldiers. The Afghan soldiers give the American soldiers fair warning of their beliefs, and the religious object is not an essential military objective, though damaging it would be better for the American soldiers because it would demoralize the Afghan soldiers. This example raises two questions:

- 1. May the American soldiers damage the religious object without violating the proportionality requirement, even if they know that the Afghan soldiers are more upset by this than by rape or torture, and is it the case that raping or torturing the Afghan soldiers would violate this requirement?*
- 2. More broadly, we see a tension between the values of cultural communities and the freedom of soldiers in a just war to do what they need to achieve victory. Are the American soldiers obliged to not engage in certain acts (rape, torture, disguise themselves as aid workers) but permitted to engage in others (damaging religiously important objects, killing religious leaders who are combatants), regardless of the values of the Afghan soldiers? Or, are they obliged to respect whatever the Afghan soldiers believe their own religious integrity consists of, as long as respecting this does not prevent essential military objectives or require the American soldiers to compromise their own deeply-held values?*

Walzer:

For me it is very important to note that in five years of fighting in Afghanistan, to the best of my knowledge, these questions have never arisen. And the reason that they haven't arisen is that rape and torture and the deliberate killing of civilians or prisoners are pretty universally taken to be wrongs—and to be more wrong than damage to religious objects. Even soldiers who rape, torture, and kill are quick to deny that they are doing those things and equally quick to accuse their enemies of doing them. The rules of *jus in bello* are not culturally specific; they can't be, since wars are fought, have always been fought, across cultural boundaries. The answer to the very last of these questions, the one about respect, is yes. But the respect at issue here doesn't follow from just war theory but from ordinary morality. We respect the "religious integrity"

of others when we can—but not when it would make us complicit, say, in the burning of Hindu widows or the sexual mutilation of African girls. However that understanding of respect is phrased, it also applies in wartime.

Topic Two:

In Just and Unjust Wars, in answer to the question “Is there a particular understanding (and then a particular distribution) of social goods that is good simply?” you write that, “That is not a question that I have addressed in this book. As a singular conception, the idea of the good does not control our arguments about justice” (312). This suggests that justice does not need to say anything about how a particular social good ought to be understood. In defense of this idea, you give the example of an Indian village where a grain of heap is divided significantly unequally between the villagers, as a result of “a long series of other inequalities... justified by customary rules and an overarching religious doctrine” (313). You argue that justice “does not rule out the inequality of the portions; it cannot require a radical redesign of the village against the shared understandings of the members. If it did, justice itself would be tyrannical” (313). I have two questions related to this:

3. *It seems that there are times when justice should have a say in determining what the social meaning of a particular social good is. First, it seems that some social meanings lead to methods of distribution that are unjust regardless of whether or not the distributions are in line with the social meanings. For instance, if the social meaning of political power entails that political power should be distributed according to ethnic background or skin color, then it seems as if the method of distributing political power would be unjust. It seems that it would be unjust even if those who lost political power on account of their race/ethnicity shared that understanding of the meaning of political power and even if distributions were made perfectly in accordance with that shared meaning. There are many real-world examples of such situations. Would you say that such practices are simply the result of short-sighted cultural bias, or is there more to it?*

There is a second reason why it seems as if justice must sometimes have a say in determining what the social meaning of a particular social good is. From your discussion of shared meanings and justice, it seems that any just society would have to have some provisions for making sure that everyone’s understanding of the meaning of a social good is taken into account in the distribution of that good. But in order for that to be possible in a society, certain shared understandings about

the meaning of legitimate social power (for lack of a better phrase) would already have to be in place: there would have to be the agreement that each person ought to be given equal influence in determining the 'shared meaning' of goods and thus in the distribution of goods. Thus, at least in this case (in the case of the meaning of 'legitimate social power'), justice could not remain neutral with regard to the content of shared meanings.

4. *If, as you say, tyranny is the opposite of complex equality and that justice requires complex equality, what would it mean for justice to be tyrannical?*

Walzer:

These questions are really one question, and I shall answer them together. Number 4 follows from a misreading of the line quoted from *Spheres*: "If it did, justice itself would be tyrannical." I mean: justice itself wouldn't be just, though it might be called that. But what about the Indian village example, which is probably the most frequently cited and criticized passage in the book? It seems obvious to my critics that distributions according to some hierarchical principle are unjust; they also believe that hierarchical principles have often been fully accepted, even by the people at the bottom ("There are many real world examples..."), and so, they conclude, shared meanings cannot play a part in our understanding of distributive justice. We require a prior theory of what is just.

My own view, expressed in the uncited paragraph that follows the one from which the quote is taken (and in subsequent pieces), is that genuinely oppressive and unjust hierarchies are not likely to be accepted by the people they oppress. The happy slave is a feature of the ideology of slaveowners, and philosophers, at a minimum, should be skeptical about that kind of happiness. Some theory of false consciousness lies behind the lack of skepticism, and behind that lies a theory of objective interests. I am sure that some people have strange ideas about their interests. And it may be true that some people, in masochistic fashion, embrace their oppression. But I don't think that can ever be the case for the larger number of oppressed people. Hence the social meanings that undergird oppressive practices are never in fact shared.

It is also important to recognize that some idealized view of hierarchy is often used to criticize the actual practices of hierarchical superiors. You claim to protect us, medieval peasants say to their feudal lords, and if you did your authority would be justifiable, but you don't protect us; instead you are the greatest threat to our well-being. You claim to mediate our relation to God, say the Christian faithful to their priests, but you serve only yourselves, at our expense. In these cases,

there might well be widespread acceptance of hierarchy in theory, but there is also criticism of and resistance to the way it works in the world.

Notice the disagreement among feminist historians and theorists about whether to stress the oppression of women or the resistance of women to their oppression. I prefer the latter, not only because it is more uplifting but also because I think it provides a more accurate picture of women's "hearts and minds."

Finally, does my argument require that each person be given equal influence in determining social meanings? If it did, then you wouldn't need anything else; the rest of the argument would be superfluous. The only just society would be some kind of radical social democracy. That is in fact my own preferred version of justice, but I don't want to impose it on everyone else, not in theory and not in practice. So it is enough if social meanings are determined by the normal social processes, in which many people participate, in different ways, over long periods of time. Take a look at my favorite example in *Spheres*—the history of the cure of souls and the cure of bodies. That is the kind of history of consciousness to which I am committed.

RESOURCES

FOR PHILOSOPHY UNDERGRADUATES

This section includes listings of journals, contests, and conferences—all of which are available to undergraduates in philosophy. If you have comments, suggestions, or questions, or if you would like to be listed here in the next issue, please contact us and we will gladly accommodate your request.

JOURNALS:

There are numerous journals, published both in print and online. The information is as recent as possible, but contact the specific journal to ensure accurate information.

Aporia: Brigham Young University. Submissions due early fall. Papers not to exceed 5,000 words. Send submissions to: Aporia, Department of Philosophy, JKHB 3196, Brigham Young University, Provo, UT 84602. Visit: <http://aporia.byu.edu/site.php?id=current>

The Bertrand Russell Society Quarterly: Edinboro University of Pennsylvania. Visit: <http://www.lehman.edu/deanhum/philosophy/BRSQ/>

Cyberphilosophy Journal: University College of the Cariboo. No posted due date. Accepts articles, books, web sites, etc. See <http://www.cariboo.bc.ca/cpj/>.

The Dialectic: University of New Hampshire. Submissions due in April. Essays (15-20 pages), short critical articles, book reviews, artwork. Send submissions to: The Dialectic, c/o Department of Philosophy, University of New Hampshire, Hamilton Smith 23, Durham, NH 03824. Visit: <http://www.unh.edu/philosophy/Programs/dialectic.htm>

Dialogue: Phi Sigma Tau (international society for philosophy). Published twice yearly. Accepts undergraduate and graduate submissions. Contact a local chapter of Phi Sigma Tau for details or write to Thomas L. Predergast, Editor, Dialogue, Department of Philosophy, Marquette University, Milwaukee WI 53233-2289. Visit: <http://www.achsnaatl.org/society.asp?society=pst>

Discourse: University of San Francisco. *Discourse* is an interdisciplinary philosophy journal featuring the work of undergraduate and graduate students. Send questions to discourse@usfca.edu. Mail submissions to: Discourse, Department of Philosophy, University of San Francisco, 2130 Fulton Street, San Francisco, CA 94117. See also <http://www.usfca.edu/philosophy/discourse/index.html>.

The Dualist: Stanford University. Submissions due February 28th. 10-30 page submissions. For more information, see <http://www.stanford.edu/group/dualist/> or contact dualist@turing.stanford.edu. Send submissions to The Dualist, Philosophy Department, Stanford University, Stanford, CA 94305.

Ephemeris: Union College. For more information, write: The Editors, Ephemeris, Department of Philosophy, Union College, Schenectady, NY 12308. Visit: <http://www.vu.union.edu/~ephemeris/>.

Episteme: Denison University. Due February 1. Maximum 4,000 words. Contact: The Editor, *Episteme*, Department of Philosophy, Denison University, Granville, Ohio 43023. Visit: <http://www.denison.edu/philosophy/episteme.html>

Ergon: University of South Carolina. Submissions accepted from undergraduate and graduate students. Visit: <http://www.cla.sc.edu/PHIL/ergon/index.html>.

Interlocutor: University of the South, Sewanee. Direct questions to Professor James Peterman at jpeterma@sewanee.edu. Send submissions to Professor James Peterman, Philosophy Department, 735 University Avenue, Sewanee, TN 37383-1000. Visit: <http://www.sewanee.edu/Philosophy/Journal/2006/current.html>

Janua Sophia: Edinboro University of Pennsylvania. Submissions and inquiries sent to Janua Sophia, c/o Dr. Corbin Fowler, Philosophy Department, Edinboro University of Pennsylvania, Edinboro, PA 16444. Visit: <http://www.edinboro.edu/cwis/philos/januasophia.html>

Meteorite: University of Michigan, Ann Arbor. Submissions due April 1. Direct questions to editors@meteorite.com and send submissions to Department of Philosophy, C/O Meteorite, University of Michigan, 435 South State Street, Ann Arbor, MI 48109-1003. Visit: <http://www.meteoritejournal.com/>.

Populas: University of California's Undergraduate Journal. Submissions limited to University of California students. Visit: <http://philosophy.ucdavis.edu/ugjournal>.

Princeton Journal of Bioethics: Princeton University. Visit <http://www.princeton.edu/~bioethic/journal/>.

Prolegomena: University of British Columbia. Visit <http://www.philosophy.ubc.ca/prolegom/> or write prolegom@hotmail.com or Prolegomena, Department of Philosophy, 1866 Main Mall, Buchanan E370, University of British Columbia, Vancouver, B.C., Canada. V6T 1Z1.

Promethius: Johns Hopkins University. *Prometheus* strives to promote both undergraduate education and research, and looks for submissions that originate from any scholarly field, as long as those submissions clearly demonstrate their applicability to philosophy. Visit <http://www.jhu.edu/prometheus/>. Write prometheusjhu@hotmail.com or Prometheus, c/o Philosophy Dept., 347 Gilman Hall, Johns Hopkins University, Baltimore, MD 21218.

Sophia: University of Victoria. Visit: <http://web.uvic.ca/philosophy/sophia/>

Stoa: Santa Barbara City College. For more information, write The Center for Philosophical Education, Santa Barbara City College, Department of Philosophy, 721 Cliff Drive, Santa Barbara, CA 93109-2394. Visit: <http://www.sbccc.edu/philosophy/website/CPE.html>.

The Yale Philosophy Review: Submissions due February 1. Visit: http://www.yale.edu/ypr/submission_guidelines.htm

CONFERENCES:

There are many undergraduate conferences, so contacting the philosophy departments of a few major schools in a particular area or researching on the web can be quite effective. The conferences below are by no means an exhaustive list.

American Philosophical Association: The APA website, <http://www.apa.udel.edu/apa/opportunities/conferences/>, contains an extensive list of conferences.

Butler Undergraduate Research Conference: Butler University. Conference in mid-April. See <http://www.butler.edu/urc/index.html> for details.

National Undergraduate Bioethics Conference: Notre Dame. Visit <http://ethicscenter.nd.edu/events/nubec.shtml> or write bioethic@nd.edu.

Ohio University Student Conference on Applied Ethics: Ohio University. Conference in late April. Contact ethics@ohio.edu. Visit <http://freud.citl.ohiou.edu/ethics/conferences.php>

Pacific University Undergraduate Philosophy Conference: Pacific University. Conference in early April. Visit <http://www.pacificu.edu/as/philosophy/conference/index.cfm> for details.

Rocky Mountain Philosophy Conference: University of Colorado at Boulder. Visit: <http://www.colorado.edu/philosophy/rmpc/rmpc.html>

Stanford Undergraduate Philosophy Conference: Stanford University. Conference in late April. Contact dualist@turing.stanford.edu for more information.

ESSAY CONTESTS:

The essay contests listed below aim at a broad range of undergraduates, but there are many other contests open to students enrolled at specific universities or interested in particular organizations. .

Elie Wiesel Essay Contest. Open to undergraduate juniors/seniors with faculty sponsor. Questions focus on current ethical issues. Due in late January. Top prize \$5,000. For more information, visit: <http://www.eliewiesel foundation.org/EthicsPrize/index.html>

ACKNOWLEDGEMENTS

The Dualist would like to thank the following contributors from Stanford University:

Philosophy Department

The ASSU Publications Board

Special Thanks:

Vera Haugh

Natasha Lee

Eve Scott

Sunny Toy

ABOUT THE DUALIST

THE DUALIST is a publication dedicated to recognizing valuable undergraduate contributions in philosophy and to providing a medium for undergraduate discourse on topics of philosophical interest. It was created by students at Stanford University in 1992 and has since featured submissions from undergraduates across North America. If you would like to receive an issue of THE DUALIST or to submit a paper, please contact us at the address below. We prefer that submissions be formatted according to APA or MLA guidelines. Papers should be submitted in paper form and by email.

Send submissions and inquiries to:

The Dualist
Philosophy Department
Stanford University
Stanford, CA 94305

To learn more about the journal, please email:

the.dualist@gmail.com

Or visit our web page at:

<http://www.stanford.edu/group/dualist>