

4-18-2006

# Anti-Individualism and Rationality

Adam J. Arico

*University of Missouri-St. Louis*, [adam.arico@umsl.edu](mailto:adam.arico@umsl.edu)

Follow this and additional works at: <http://irl.umsl.edu/thesis>

---

## Recommended Citation

Arico, Adam J., "Anti-Individualism and Rationality" (2006). *Theses*. 199.  
<http://irl.umsl.edu/thesis/199>

This Thesis is brought to you for free and open access by the Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Theses by an authorized administrator of IRL @ UMSL. For more information, please contact [marvinh@umsl.edu](mailto:marvinh@umsl.edu).

Anti-Individualism and Rationality

Adam Arico  
Master's Thesis  
Department of Philosophy  
University of Missouri – St. Louis  
18 APR 2006

ABSTRACT: Paul Boghossian has argued that *de re* ascriptions of mental content are unfit for assessments of rationality.<sup>1</sup> An individual, he suggests, can comport with the principles of logic and rationality if and only if she is able to introspectively recognize the content of her thoughts and beliefs without relying on empirical evidence. Jessica Brown argues, first, that Boghossian's notion of rationality is unrealistic and incompatible with modern experimental data, and second, that his principles of transparency of mental content are unreasonable and, ultimately, self-defeating.<sup>2</sup> In defending Boghossian's notion of rationality, and subsequently his transparency principles, this paper attempts to respond to Brown's arguments by distinguishing between two notions of rationality (weak and strong) and between two aspects of mental content (psychological and social). Acknowledging the weak notion of rationality will preserve the ordinary conception of typical subjects as rational beings, while introducing the distinction between psychological content and social content resolves the contradiction between the principles of transparency in Mates case circumstances.

## Introduction

In a chapter of her recent book, *Anti-Individualism and Knowledge*, Jessica Brown responds to the criticism, put forth by Paul Boghossian, that anti-individualism is incompatible with our ordinary concept of rationality. Boghossian's criticism is that, since anti-individualism denies epistemic transparency of mental content, it undermines an otherwise rational subject's ability to behave rationally (i.e., denying transparency disposes agents to holding contradictory beliefs and/or making invalid inferences). Boghossian argues that any theory of mental content, if it is to comport with even a minimal notion of rationality, must conform to the following two principles:

---

<sup>1</sup> Boghossian, (1994).

<sup>2</sup> Brown, (2004).

***Transparency of sameness of content:** If two of a thinker's token thoughts possess the same content, then the thinker must be able to know a priori that they do*  
***Transparency of difference of content:** If two of a thinker's token thoughts possess distinct contents, then the thinker must be able to know a priori that they do.*<sup>3</sup>

Brown, however, argues that the notion of rationality espoused by Boghossian is unrealistically demanding and inconsistent with the behavior displayed by subjects in recent psychological experiments. Moreover, she argues that Boghossian's principles of transparency fail regardless of the truth of anti-individualism, that there are situations in which the principles of epistemic transparency of mental content are self-defeating. Thus, she says, it cannot be an objection against anti-individualism that it undermines the transparency of a subject's mental content and, thereby, Boghossian's notion of rationality. This paper examines and responds to both arguments, first by pointing to an alternate notion of rationality also espoused by Boghossian—one that is perfectly consistent with the psychological data. Second, it attempts to show how, even in the special circumstances described by Brown, transparency of mental content can be preserved by distinguishing between the psychological content of thought and the social content of an expression. Since it is important to understand the background criticisms to which Brown is responding before presenting her arguments against Boghossian, the following section presents the thought-experimental intuitions that motivate Boghossian's principles of transparency.

## **1. Anti-Individualism and Irrationality**

Boghossian argues that an anti-individualist conception of content undermines our ordinary conception of rationality, because it precludes a subject's a priori ability to make simple valid inferences, to avoid holding contradictory beliefs, and to avoid making invalid inferences.

---

<sup>3</sup> Cf. Boghossian, p. 42.

Without epistemic transparency of mental content, he says, “our conceptions of rationality and of rational explanation yield absurd results.”<sup>4</sup> For each principle of transparency there is a paradigm thought experiment<sup>5</sup> meant to show how rejecting a priori access to mental content prevents a subject from acting in a way that can be deemed rational. First, I present a thought experiment provided by Brian Loar, meant to elucidate the need for transparency of sameness of content; second, to support the need for transparency of difference of content, I will describe a situation that arises from Hilary Putnam’s famous “Twin Earth” thought experiment.

### **1.1 Sameness of Content**

In support of the idea that rational behavior requires the transparency of sameness of content, consider the following example from Loar:

Suppose that Paul, an English speaker, has been raised by a French nanny in a sheltered way. She speaks English with Paul, but amuses herself by referring to the cats around them as “chats” (she says “shahs,” pronouncing the “s”) and never as “cats.” Paul acquires thereby a perfectly good recognitional acquaintance with cats and many beliefs about them, but he does not know that in English they are properly called “cats.” Suppose he forms the belief he would express as “All chats have tails”; it seems [by an anti-individualist description of his mental content] we are then justified in asserting the Paul believes all cats have tails... As it happens, he occasionally sees his parents, who speak of animals called “cats.” Because no cats are ever present, nor any pictures of cats, Paul does not realize that cats are his familiar “chats.” Now Paul’s parents tell him various things about cats, in particular that they all have tails. On this basis it is again true of Paul that he believes that all cats have tails.<sup>6</sup>

If we take an anti-individualist position we have to admit that Paul can come to hold two beliefs with the same content, namely that cats have tails. But, as both Boghossian and Loar point out, “intuitively...it is quite clear that Paul will not be able to tell *a priori* that the belief he expresses

---

<sup>4</sup> Boghossian, p. 39.

<sup>5</sup> That is to say, there is *at least* but not *only* one thought experiment utilized for the purpose of demonstrating that lack of transparency entails lack of rationality.

<sup>6</sup> Loar, (1988), p. 184.

with ‘All chats have tails’ is the same belief that he expresses with ‘All cats have tails’.”<sup>7</sup> This example clearly illustrates how anti-individualism interferes with the transparency of sameness of mental content.

From this interference with transparency, Boghossian contends, we can see how anti-individualism potentially leads to irrationality, since Paul, lacking a priori access to his thoughts, becomes vulnerable to holding beliefs with contradictory content. If, for instance, Paul’s parents talked all and only about Manx cats, then one of the various facts that he would have learned about “cats” is that some of them do not have tails. From this it would follow that Paul believes both “Some cats do not have tails” and (from the nanny’s lessons about all “chats” having tails) “All cats have tails.” But these two beliefs are blatantly contradictory. Thus, if the contents of his thoughts are determined externally, Paul will unwittingly hold beliefs with contradictory content.

Both Boghossian and Loar appeal to our intuitions in utilizing this thought experiment. It seems intuitive, they say, that Paul has distinct concepts with distinct content, since his beliefs about “chats” serve different mental functions and interact with other beliefs in markedly different ways than his beliefs about “cats.” Since they serve different mental functions and interact in different ways with other beliefs, it seems natural to say that Paul’s content expressed by “chats” is different than his content expressed by “cats.” Paul’s parents may very well have instilled in him the axiom “never bring a cat to school,” but it is intuitively difficult to accuse Paul of being either disobedient or irrational for bringing his favorite “chat” to show-and-tell. All of these intuitions, though, would likely change if we understood Paul as recognizing that his beliefs about “cats” and “chats” had identical content. That is, it seems that we can only fairly

---

<sup>7</sup> Boghossian, p. 37.

assess Paul's rationality if the identical content of his beliefs is available to him upon introspection, which just is the principle of sameness of mental content.<sup>8</sup>

## **1.2 Difference of Content**

Boghossian also points to instances in which irrationality follows from the fact that “two tokens of a thinker's language of thought belonging to the *same* syntactic type have *distinct* meanings, but the thinker is [under an anti-individualist interpretation] not in a position to know *a priori* that they do.”<sup>9</sup> Such instances arise when we consider what he calls ‘one-world traveling cases.’ Hilary Putnam's Twin-Earth thought experiment is supposed to illustrate the necessity of individuating mental contents by the external objects to which they refer.<sup>10</sup> As one version of the story goes,

There is a world, call it “Twin Earth,” which is identical to this world in every way save one: on Twin Earth, the stuff in rivers and lakes and faucets is composed of a substance superficially indistinguishable from water, but with chemical structure XYZ instead of H<sub>2</sub>O. Without consent or knowledge, the story continues, an earthling named Oscar is seamlessly transported to Twin Earth (in his sleep, perhaps), where he forms beliefs and utters propositions about the stuff in twin lakes, twin rivers, and twin faucets. Anti-individualism suggests that, since there is no H<sub>2</sub>O in Oscar's environment, ‘water’ now picks out instances of XYZ, since his environment individuates the content of Oscar's beliefs and utterances.<sup>11</sup> Oscar, by most accounts (including Brown's), also uses ‘water’ to talk and think about aquatic interactions that occurred on Earth, instances of interactions with H<sub>2</sub>O. According to most versions of anti-individualism, then,

---

<sup>8</sup> In an earlier chapter, Brown argues that her version of anti-individualism, unlike other popular versions, can successfully accommodate the principle of transparency of sameness of content by utilizing a ‘Fregean sense,’ in which sense is understood as “a way of thinking about an object that would not be available in the absence of the object.” (Brown, p. 21) Paul may think about cats via a ‘cats’ sense, and he may think about cats via a ‘cats’ sense; since these are two different ways of thinking about cats, they are two different senses with two different contents. Thus, utilizing her ‘Fregean anti-individualism,’ Brown believes she can avoid attributing logically contradictory beliefs to Paul, as well as any other cases of identical referential content. My suspicion is that this position appears more like a two-factor or dual-content theory (one part ‘object’, one part ‘way of thinking about object’) than a purely anti-individualist position. I do not, however, pursue this suspicion any further here.

<sup>9</sup> Boghossian, p. 37.

<sup>10</sup> Cf., Putnam (1975).

<sup>11</sup> Some anti-individualists hold that the change in reference of ‘water’ occurs immediately; others who hold that the change occurs only after a suitable amount of time; still others who hold that ‘water’ comes to possess a composite meaning, picking out both H<sub>2</sub>O and XYZ. Brown, however, dismisses the last theory as one that “would not be accepted by many anti-individualists” and concludes that most anti-individualists “accept that, as a result of a slow switch, a subject may have two concepts that she expresses by a single term.” (Brown, p. 185).

Oscar has two indistinguishable concepts: one ('water') that picks out H<sub>2</sub>O and another ('twater') that picks out XYZ. However, Oscar is ignorant of which concept he has in mind when thinking about water.

According to Boghossian, this inability to distinguish between the content of his thoughts could lead Oscar to behave irrationally, since it opens him up to the possibility of making invalid inferences. Oscar may have formed a general belief rooted in an Earthly experience that looks something like "water quenches my thirst," whose content is 'water.' After a particularly grueling workout on Twin Earth, Oscar eyes a bottle of 'twater' and concludes (based on his general belief about 'water'), "This will quench my thirst." Under the anti-individualist description of things, the 'water' content of his general belief differs from the 'water' content of the particular bottle. Thus, Oscar's inference to the conclusion that the bottle of 'twater' will quench his thirst is invalid, though he sees it as a perfectly valid inference. The invalidity of the inference (and thus, Oscar's irrationality), Boghossian says, is inevitable so long as the content of Oscar's concepts remains unavailable to introspection. That is, Oscar is irrational because he lacks transparency of mental content.

These two thought experiments are meant to convey the counter-intuitive consequences of an anti-individualist semantic theory. A brief sketch of Boghossian's argument looks something like this:

1. Being rational is a matter of making valid inferences and avoiding contradictions.
2. In order to make valid inferences and avoid contradictions, subjects must have a priori, introspective access to the content of their thoughts.
3. If we assume an anti-individualist semantics, then subjects do not have a priori, introspective access to their thoughts (as demonstrated by the Paul and Oscar).
4. Thus, if we assume an anti-individualist semantics, subjects are not rational.

Brown denies both Boghossian's notion of rationality (1) and his belief that anti-individualism is to blame for subjects failing to have a priori, introspective access to the content



of their thoughts (3). The following section presents and, in turn, responds to these two arguments.

## **2. Challenging Boghossian**

Brown says that her proposal to reject Boghossian's notion of rationality "is also motivated by considerations that are independent of the truth of anti-individualism. There is much evidence that subjects do not always conform their thoughts to the laws of logic a priori."<sup>12</sup> She gives two counter-examples to support her claim that transparency, when combined with Boghossian's notion of rationality, fails independently of the veracity of any particular semantic theory. First, she argues that subjects in psychological experiments typically fail to display the kind of rationality described by Boghossian, even in circumstances where transparency of mental content is not involved or required. Second, she points to certain situations—namely, what she calls Mates cases—in which transparency of mental content unavoidably fails, whether we assume an anti-individualist position or not. I want to reply to her first argument by calling attention to a second notion of rationality, a weaker notion (entertained by Boghossian, but not mentioned by Brown), which might be capable of dealing with the problematic experimental data. In response to the second counter-example, I attempt to demonstrate that, even in the Mates cases described by Brown, transparency can be maintained.

### **2.1 Wason's Selection Task**

Brown begins her argument against Boghossian's notion of rationality with the observation that "subjects sometimes make mistakes in their reasoning, or have inconsistent beliefs, when only a little logical acumen is required to reveal the mistake."<sup>13</sup> The failure to avoid or notice such inconsistencies is commonplace, she says, even among highly intelligent

---

<sup>12</sup> Ibid., p. 185.

<sup>13</sup> Ibid., p. 185.

and educated persons. Take, for instance, the regularity of less than logical behavior displayed by typically rational subjects in Wason's selection task:

Subjects are presented with four cards, each with only one side visible; they show 'A', 'K', '4', and '7' respectively. Subjects are asked which cards they need to turn over in order to test the truth of the rule: 'If a card has a vowel on one side, then it has an even number on the other.'

Most subjects respond, quite correctly, that the 'A' card needs to be turned over, but the surprising and relevant discovery that Wason (and numerous subsequent experimenters) made was that many subjects also respond, quite incorrectly, that the '4' card also needs to be turned over. The appropriate method of testing the rule's validity would be to turn over the '7' card, as its possessing a vowel would demonstrate a failure of the rule. Brown argues that this experiment illustrates that the vast majority of subjects seem neither able nor disposed to conform to the laws of logic *a priori*, contrary to Boghossian's claims about rationality. *A priori* conformity to the laws of logic is just the kind of rationality that the principles of transparency are designed to ensure, but this kind of rationality seems to be more than what most people are capable of. If *a priori* conformity to the laws of logic is beyond the range of typical human cognition, then, Brown argues, any notion of rationality that requires such conformity is unrealistic and should be dismissed.

It is not especially controversial to interpret the participants in the selection task as failing to behave as if they were *disposed* to adhere to the laws of logic, *a priori* or otherwise. If this disposition were essential to Boghossian's notion of rationality, then Wason's evidence would surely be sufficient to undermine it. It would follow, as Brown argues, that any principles governing propositional content aimed at maintaining such a notion of rationality would be unrealistically demanding. It would then follow that, if the principles governing propositional content were unrealistic in their demands, then it couldn't possibly be an argument against anti-

individualism that it fails to comport with those principles. In the next section, I will point to a second notion of rationality mentioned in Boghossian's paper, a notion that does not seem to conflict with the data from selection task experiments.

## **2.2 Weak versus Strong Rationality**

If we take Brown at face value, then it seems that Boghossian's notion of rationality is overly demanding and unrealistic in its constraints. There are, however, two notions of rationality that Boghossian depicts in setting the stage for the principles of transparency of content: one I refer to as *weak rationality*, the other as *strong rationality*. Though I suspect both of them are equally consistent with the principles of transparency, they are not equally resilient to the charges set forth in regard to the selection task—namely, that Boghossian's rationality conflicts with the behavior displayed in numerous psychological experiments. Brown's argument that Boghossian's rationality is inconsistent with experimental data turns on what should be seen as *strong rationality*, which is suggested by the following passage:

What does a person have to do to count as a good reasoner? Clearly, it is not at all a question of knowing empirical facts, of having lots of justified true beliefs about the external world. *Rather, it is a matter of being able, and of being disposed, to make one's thoughts conform to the principles of logic on an a priori basis.*<sup>14</sup>[my emphasis]

However, we need not understand Boghossian as endorsing only *strong rationality*. In fact, I take him to be equally committed to a significantly weaker notion of rationality, wherein an agent is rational insofar as she is able, given sufficient opportunity to reflect on the matter, to avoid obvious violations of the laws of logic. In a portion of the paragraph quoted by Brown (in which she utilizes ellipsis to omit the weaker claim from the block quoted in her book), Boghossian admits:

---

<sup>14</sup> Boghossian, p 42.

We may, if we wish, put the matters in a far less committal way: let's say that being *minimally* rational is a matter of being able to *avoid* obvious violations of the principles of logic, given enough time to reflect on the matter and so on.<sup>15</sup>

Under the *strong* notion of rationality, agents are not only *able*, they are *disposed* to conform to the laws of logic a priori, which suggests that humans, in general, tend to behave in ways that comport with the laws of logic. This idea resembles what some social scientists call “hyperrationality,” which is the normative or idealized conception of humans as maximally rational agents and ideal decision makers. Similarly, economic utility theory depicts humans as optimizing agents that generally make the most rational decision and maximize their welfare in any situation when given sufficient relevant information.<sup>16</sup>

The *weak* notion of rationality, however, says only that humans are generally *able* to be rational, in that they have the ability to avoid *obvious* violations of logical principles. The *weak* notion is much more consistent with today's commonly accepted beliefs regarding human rationality, e.g., Herbert Simon's theory of bounded rationality.<sup>17</sup> Bounded rationality is the idea that human rationality is limited by factors, such as emotions, superstitions, lack of time and memory space, etc. The *weak* notion of rationality, like the notion of bounded rationality, admits that there are factors that limit the rationality of human behavior; however, it also claims that, given enough time and opportunity, agents can overcome those factors to determine whether their behavior conflicts with basic logical principles.

---

<sup>15</sup> Boghossian, p 42.

<sup>16</sup> This model is sometimes taken as a normative description of the behavior humans should strive for: “Political Economy considers mankind as occupied solely in acquiring and consuming wealth; and aims at showing what is the course of action into which mankind, living in a state of society, would be impelled, if that motive... were absolute ruler of all their actions.” [Mill (1836), p. 53].

On the other hand, it is sometimes used as a descriptive theory of rationality, useful for predicting human behavior: “[Positive Economics’] task is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances.” [Friedman (1953), p. 4].

<sup>17</sup> Cf., Simon (1957).

Nothing in this weaker notion of rationality is *prima facie* inconsistent with the behavior displayed in the selection task. Though subjects failed to utilize the ideal logical principles that would most fully test the truth of the experimental rule, since they do not seem to be acting in direct violation of any logical principles, they do not fail to be rational insofar as we utilize the *weak* notion of rationality. As such, they appear to satisfy the conditions of rationality, if rationality consists in the *weak* rather than *strong* notion.

All that can be said of the selection task participants is that they are not *maximally* rational agents (i.e., “hyperrational” optimizers), and that they fail to be rational only on the *strong* notion of rationality. They do not appear to be relying on obviously contradictory lines of reasoning, so despite their general failure to reach ideally rational conclusions, they nonetheless qualify as rational beings on the *weak* notion. It is essential, however, to Brown’s argument that the selection task participants act contrary to Boghossian’s notion of rationality. If they do not fail to count as rational beings, then Boghossian’s principles of transparency cannot be charged with being based on an unrealistic notion of rationality. Because Brown’s argument hinges on the *strong* notion of rationality, distinguishing between it and *weak* rationality preserves the both Boghossian’s notion of rationality and his principles of transparency. The next section explores whether a similar distinction can rescue transparency from Brown’s second attack.

### **2.3 Mates Cases**

The second argument against Boghossian is intended to show that, regardless of the truth of anti-individualism, there are situations in which the principles of transparency of content are self-defeating. That is, there are times in which one rational individual’s ability to recognize a priori sameness (or difference) of thought content precludes another rational individual’s ability to recognize a priori difference (or sameness) of thought content. If Brown can successfully

demonstrate that in at least some circumstances transparency fails regardless of the truth of anti-individualism, she will rebut the objection to anti-individualism that it should be rejected to preserve epistemic transparency of mental content.

Brown's task, then, is to describe at least one situation in which a failure of transparency necessarily occurs, in which it is impossible for at least one person to have a priori access to the similarity or difference of her thought contents. She thinks such a situation arises in certain instances, namely those instances that have come to be known as Mates cases.<sup>18</sup> Take two sentences:

- (i) No one doubts that whoever believes that Mary is a physician believes that Mary is a physician.
- (ii) No one doubts that whoever believes that Mary is a physician believes that Mary is a doctor.

There is some disagreement amongst philosophers as to whether (i) and (ii) express the same proposition; that is, whether substituting 'doctor' for 'physician' in (ii) preserves the content in (i). Mates himself argued that it is possible for someone to accept (i) yet doubt (ii); after all, (i) is an analytic truth that cannot fail to be true, while all that is necessary for (ii)'s veracity to fail is for one individual to doubt that whoever believes that Mary is a physician believes that Mary is a doctor. Others claim that it is impossible to accept (i) yet deny (ii).<sup>19</sup> Brown utilizes this debate to attack transparency:

This disagreement about whether the sentences (i) and (ii) express the same proposition can be used to argue that there is a pair of thoughts that each philosopher thinks although they disagree about whether these thoughts have the

---

<sup>18</sup> The Mates case was first presented in Benson Mates' 1950 paper "Synonymy."

<sup>19</sup> The origin of the disagreement arose between Mates and Church on whether Carnap's explication of synonymy as *intentional isomorphism* (Carnap argued that two words were synonymous if they could be substituted *salve veritate* in any sentence) was satisfactory. Mates believed it to be obvious that (i) could be true while (ii) could be false, using this 'obvious' truth as a counter-example to Carnap. Church (1954) argued against Mates that, since anyone who could be described as believing that Mary was a physician could also be described as believing that Mary was a doctor, (ii) also had to be true.

same or different content.<sup>20</sup>

The argument develops into the following dilemma: either the two sentences express the same content or they express different content. Since some believe that the two express identical content, while others believe that the two express different content, we inevitably encounter a failure of transparency. If the two sentences express identical content, then those who believe they express different content are incapable of recognizing the sameness of content a priori; but if the two sentences express different content, then those who believe they express the same content are incapable of recognizing the difference of content a priori. Whichever horn of the dilemma we grasp, we are left with a failure of transparency. Importantly, Brown thinks this holds true regardless of the veracity of anti-individualism. I do not think, however, that Brown sufficiently explored the viability of transparency in Mates cases under an individualist position. What follows is an attempt at just such an exploration.

#### **2.4 Psychological Content**

If the Mates case truly poses a problem for transparency of content *regardless of the truth of anti-individualism*, then transparency should fail even if we distinguish between the social usage of a term and the individual's use of that term. It may be that the content of thought and the content of linguistic expression are not identical to each other, that there can be a difference between the psychological content of a thought associated with a given expression and the social content of that expression in the public language. I develop this possibility more fully below, but first allow me to explain the general structure of my argument.

It might be obvious, already, how this exploration will proceed: describe the Mates case scenario in an individualist framework, and demonstrate that transparency doesn't fail. And it might seem that I am just guilty of begging the question and stacking the chips in my favor.

---

<sup>20</sup> Brown, p. 186.

Remember, though, that Brown was ostensibly demonstrating that transparency in Mates cases is self-defeating *even if anti-individualism were not true*, from which it would follow that failure of transparency cannot be an objection against anti-individualism. Since, I suggest, her presentation of the Mates case was not sufficiently disentangled from anti-individualism, it falls to me to demonstrate that transparency in Mates cases is possible when we assume anti-individualism is not true. From this demonstration, it would then follow that transparency does not fail regardless of anti-individualism.

If we can reasonably distinguish between psychological and social content, then Brown's putative counter-example may not illustrate what it sets out to illustrate: namely that there are cases in which transparency is self-defeating. If Brown's counter-example is undermined, then Boghossian's principles of transparency, and his corresponding notion of rationality, will retain their plausibility and will remain a problem for anti-individualism. It is important, then, to see how the Mates case works when the content of thought (psychological content) is distinguished from the content of expression in public language (social content). First, however, I need to clarify what I mean by *psychological content* and *social content*.

Leaning heavily on Loar's<sup>21</sup> work, I take psychological content to be the way a person conceives of an object, sentence, term, etc. Psychological content is individuated by the conceptual or cognitive role it plays in the individual's psychology (i.e., how it interacts with and affect other beliefs, attitudes, perceptions, etc.). In the case of Paul's cats, "chats" plays a significantly different role in his belief-formations and attitudes about cats than that played by "cats." Most obviously, in the case where Paul's parents spoke only of Manx cats, "chats" interacts directly with Paul's beliefs about animals with tails, whereas "cats" does not. "Chats"

---

<sup>21</sup> Op. cit.



falls within his mental category of ‘things I can take to show-and-tell’ while “cats” does not.

These differences in conceptual/cognitive role, I suggest, entail different psychological contents.

On the other hand, by social content I mean whatever is captured by that-clauses in descriptions of meaning. These are the *de re* ascriptions which are important for capturing the references of thoughts. As Loar suggests, “that-clauses capture how a belief would be expressed by exhibiting something that is equivalent in *social content* (as we might say) to what the subject would utter, given his deference to the usage of his linguistic community.”<sup>22</sup> Whatever Paul’s “chats” thoughts and “cats” thoughts commonly refer to, that is the social content of Paul’s thoughts (namely, in this case, they both have as their extensions cats). What follows is an implementation of the psychological content/social content distinction in regard to the Mates case.<sup>23</sup>

In her discussion of the Mates case, Brown utilizes Kripke as representative of the set that believes (i) and (ii) express the same thought content, while she takes Burge to represent those who believe (i) and (ii) express different thought contents. In the name of continuity, I will do the same.

Supposing, then, that there is a purely psychological component of Kripke’s thought content, as well as a social content of the expression he would associate with that thought content, we have reason to deny that transparency is self-defeating. That is, part of the thought Kripke expresses by uttering (i) is determined by the role it plays in his psychology and by how it interacts with other beliefs, attitudes, perceptions, etc.; likewise for Burge. It is possible and, I think, quite likely that the psychological content Kripke might express by (i) is distinct from the psychological content Burge would express by (i); and it is equally possible that the

---

<sup>22</sup> Loar, p. 190.

<sup>23</sup> How the two types of content are related (e.g., how they impact and effect one another) is not a matter I address in this paper.

psychological content Kripke might express by (ii) is not identical to the psychological content Burge would express by (ii). However, if we consider psychological contents as being distinct from social contents, it would follow that two psychological contents associated with a single linguistic expression need not be identical to each other. Rather than just the two social contents expressed by (i) and (ii)—in this case denoted by (i\*) and (ii\*)—there could be up to four psychological contents:

- (Ki\*) the psychological content Kripke associates with (i),
- (Bi\*) the psychological content Burge associates with (i),
- (Kii\*) the psychological content Kripke associates with (ii), and
- (Bii\*) the psychological content Burge associates with (ii).

That is all to say that there could be as many as four unique conceptual/cognitive roles played by the concepts expressed in (i) and (ii). I will call this the four-thought model, which I suggest de-fangs Brown's argument and preserves the plausibility of the principles of transparency.

Brown argued that either (i) and (ii) express the same content or they do not, and since Kripke and Burge disagree about just that, one of them will always lack transparency of content; that is, necessarily, one of them will be unable to realize a priori the sameness or difference of his thought content. However, this argument operated on a limited conception of the content under consideration. Once we distinguish between the psychological and social contents, we can see how both thinkers can maintain transparency. Regardless of whether (i\*) and (ii\*) are identical social content, the psychological content Kripke and Burge associate with (i) and (ii) arguably can differ for each thinker and, just as importantly, between thinkers. Within the framework of psychological content vs. social content, there is no reason to believe that the psychological content (Ki\*) that Kripke associates with the public expression (i) need be identical to the psychological content (Bi\*) that Burge associates with the public expression (i).

Similarly, the psychological content ( $K_{ii}^*$ ) that Kripke associates with the public expression (ii) need not be identical to the psychological content ( $B_{ii}^*$ ) that Burge associates with the public expression (ii).

Also, whether  $(K_i^*)=(K_{ii}^*)$  will be a function of the conceptual/cognitive role(s) each thought plays in Kripke's psychology. That is, they are identical only insofar as they interact in identical ways with identical beliefs, attitudes, perceptions, etc. The converse holds true for  $(B_i^*)$  and  $(B_{ii}^*)$ : if they are different, they are so in virtue of interacting with different beliefs, attitudes, perceptions, etc.; or in virtue of interacting in different ways with the same beliefs, attitudes, perceptions, etc. If we apply the principles of transparency strictly to the psychological contents of thought, then it is possible that both Kripke and Burge can have a priori introspective access to their own thoughts without defeating the other's transparency.

The Mates case challenges transparency if and only if it selectively considers just the social content of thought. The argument against transparency is that, if  $(i^*)$  and  $(ii^*)$  are identical as Kripke thinks, then Burge lacks transparency of sameness, and if  $(i^*)$  and  $(ii^*)$  are different as Burge thinks, then Kripke lacks transparency of difference. However, even if Kripke is "correct" and the social contents of (i) and (ii) were identical, it still does not follow that the psychological contents,  $(B_i^*)$  and  $(B_{ii}^*)$ , would have to be the same. The principles of transparency, in the four-thought model, merely suggest that if the psychological contents of Kripke's thoughts are identical then he can know so a priori. Kripke need merely consider  $(K_i^*)$  and  $(K_{ii}^*)$  to determine if they are the same. That the social contents are identical, though, mandates nothing concerning the identity (or non-identity) between the psychological content of Burge's thoughts. We can maintain Boghossian's principle of transparency of sameness of content, but require only this: if  $(K_i^*)$  and  $(K_{ii}^*)$  are identical then *Kripke* must be able to know a priori that they are.

If, on the other hand, Burge is “correct” and the social contents of (i) and (ii) were different, it would also not follow that the psychological contents,  $(K_i^*)$  and  $(K_{ii}^*)$ , would have to be different. Neither the fact that Burge’s psychological contents are different nor the possibility that the social contents of (i) and (ii) are different would require anything of Kripke’s thought contents. Boghossian’s principle of transparency of difference of content merely requires that *Burge* be able to know a priori that  $(B_i^*)$  and  $(B_{ii}^*)$  are different.

When Brown presented the Mates case, she described the disagreement between Burge and Kripke as a disagreement of whether (i) and (ii) expressed the same proposition. Under Brown’s presupposition, this was simply a question of whether or not the contents of the two expressions were identical. I have suggested that this description leads to failure of transparency because it does not distinguish between the psychological content of thoughts and the social content. However, on the four-thought model, they can both be right (or wrong) without entailing anything about the other: Kripke can believe that  $(K_i^*)=(K_{ii}^*)$ , without contradicting Burge’s belief that  $(B_i^*)\neq(B_{ii}^*)$ . It remains, however, for me to show how we should understand the disagreement between Burge and Kripke under the four-thought model.

The short answer is that neither Kripke nor Burge consider the psychological/social content distinction. They are both anti-individualistic, in that they both deny that there is a content of thought wholly determined by a conceptual/cognitive role. Thus their disagreement is a matter of whether or not the social contents of (i) and (ii) are identical.

## **Conclusion**

Brown presented two arguments aimed at challenging the principles of epistemic transparency of mental content. First, she argued that the rationality motivating the principles was an unrealistic notion of rationality that conflicted with empirical data from psychological

experiments. Second, she attempted to show that, in certain situations, the principles of transparency were unavoidably self-defeating. As such, she argued, we should reject transparency-based rationality both because subjects typically fail to demonstrate an ability to follow the laws of logic a priori and because, regardless of the truth of anti-individualism, Boghossian's principles of transparency are potentially self-defeating. I have argued in response that her counter-examples do not adequately demonstrate either that subjects are incapable of following the laws of logic a priori or that Boghossian's principles of rationality necessarily fail in Mates cases. I did so by first drawing attention to a useful distinction between *strong* rationality and *weak* rationality, and then by presenting an alternative understanding of the Mates case scenario. This alternative understanding distinguishes between the psychological content of a thought and the social content. In short, I have defended Boghossian's principles of transparency, as well as his notion of rationality, from Brown's attacks.

### Works Cited

- Boghossian, Paul, (1994). "The Transparency of Mental Content." *Philosophical Perspectives* 8, pp. 33-50.
- Brown, Jessica, (2004). Anti-Individualism and Knowledge. MIT Press.
- Burge, Tyler, (1978). "Belief and Synonymy" in *The Journal of Philosophy*, Vol 75 number 3, 119-138.
- Church, Alonzo, (1954). "Intentional Isomorphism and Identity of Belief" in *Philosophical Studies*, Vol. V number 5, 65-73.
- Friedman, Milton, (1953). "The Methodology of Positive Economics" in *Essays in Positive Economics*. University of Chicago Press, pp. 3-43.
- Kripke, Saul, (1979). "A Puzzle About Belief" in *Meaning and Use*, A. Margalit (ed), Dordrecht Reidel, Netherlands, 239-283.
- Loar, Brian, (1988). "Social Content and Psychological Content" in *Thought and Content*, Robert Grimm and Daniel Merrill (eds.), University of Arizona Press, pp. 99-110.
- Mates, Benson, (1950). "Synonymy" in *Meaning and Interpretation*, University of California, Publications in Philosophy, 25, 201-226
- Mill, John Stuart, (1836). "On the Definition and Method of Political Economy" excerpted from "On the Definition of Political Economy and the Method of Investigation Proper to It." Reprinted in *Essays on Some Unsettled Questions of Political Economy* (1844), 3d ed., London: Longmans Green & Co., 1877, pp. 120-64. Reprinted in The Philosophy of Economics: An Anthology. Daniel Hausman, (ed.), Cambridge University Press, 1994, 52-58

Putnam, Hilary, (1975). "The Meaning of 'Meaning'." In *Philosophical Papers, Vol. 2: Mind, Language, and Reality*. Cambridge University Press.

Simon, Herbert, (1957). Models of Man: Social and Rational. New York: John Wiley and Sons, Inc.