

# UC Santa Barbara

## UC Santa Barbara Electronic Theses and Dissertations

### Title

Unified Agency, Rational Lies, and the Murderers at the Door

### Permalink

<https://escholarship.org/uc/item/3zk7g6kc>

### Author

Batoon, Arnel Blake Escobal

### Publication Date

2024

Peer reviewed|Thesis/dissertation

University of California  
Santa Barbara

**Unified Agency, Rational Lies, and the Murderers at  
the Door**

A dissertation submitted in partial satisfaction  
of the requirements for the degree

Doctor of Philosophy  
in  
Philosophy

by

Arnel Blake Escobal Batoon

Committee in charge:

Professor Kevyn Falvey, Chair  
Professor Peter Graham  
Professor Michael Rescorla  
Professor Nathan Salmon

March 2024

The Dissertation of Arnel Blake Escobal Batoon is approved.

---

Professor Peter Graham

---

Professor Michael Rescorla

---

Professor Nathan Salmon

---

Professor Kevyn Falvey, Committee Chair

March 2024

Unified Agency, Rational Lies, and the Murderers at the Door

Copyright © 2024

by

Arnel Blake Escobal Batoon

In loving memory of my father, Arnel Batoon.

## Acknowledgements

Writing this dissertation has been long and difficult. So many people have helped me along this journey. I want to thank my committee chair, Kevyn T. Falvey. In so many ways, Kevyn has provided for me the model of how I want to be as an academic. They showed me, through their own example, how to balance compassion and patience with academic rigor. I aspire to carry on their tradition of exemplary mentorship.

I also want to individually thank each of my committee members. I want to thank Michael Rescorla for his mentorship in my early development as a philosophy Ph.D. It's because of him that I am unafraid to delve into the difficult waters of epistemology and philosophy of psychology. I want to thank Peter Graham for being one of the most fun teachers I have had in my late graduate career. He reminded me how epistemology could be fun, even when it's difficult. I want to thank Nathan Salmon for giving me one of the most positive "meet-your-heroes" experiences a student could ever ask for.

Beyond my committee, I want to thank my mentors during my undergraduate and M.A. years at Brandeis University. Richard Gaskins, through his first year seminar on law and authority, introduced me to the depth of the philosophy of law. His advising as my first year mentor helped me navigate a difficult freshman year. Beri Marušić, with his introductory philosophy course, solidified my interest in philosophy and established epistemology as my philosophical home. His patient and thoughtful guidance throughout my undergraduate, and then M.A. career, fostered in me the sense of curiosity that animates me as a philosopher. Palle Yourgrau, throughout my undergraduate and M.A. career, showed me how to follow my own philosophical convictions, even if they bring me to surprising and mysterious places. Jennifer Smalligan Marušić, during my M.A. career, helped me explore new spaces not only in epistemology but also in the history of philosophy and philosophy of science. Marion Smiley did the same for me with ethics

and political philosophy. Their combined mentorship and encouragement through those years have influenced my philosophical voice in ways that are subtle, but profound.

Beyond my mentors and committee, I want to thank Dan Korman and Matt Hanser. The doctoral process is all too often a source of anxiety. During their terms as graduate advisors at UCSB, their patient encouragement helped me move past those anxieties.

I want to thank all of my teachers throughout my time as a philosophy graduate student. I particularly want to thank Alan Berger, Tyler Burge, Sonny Elizondo, Katie Elliott, Eli Hirsch, Thomas Holden, Sophia Malamud, Kate Moran, Sherri Roush, Jerry Samet, Eugene Sheppard, Brett Sherman, Andreas Teuber, and Aaron Zimmerman. Their teaching helped me see philosophy as a living discipline with fruitful interactions with other areas of study.

I want to thank all my fellow graduate students during my times at Brandeis and UCSB. It's really true when they say that you learn the most through interacting with your peers. This true, even for the friends I made who weren't in my home departments. Even more than that, I cherish their friendship and their support beyond the graduate school experience. On this score, I especially want to thank Michael Augustin, Katie Bainbridge, Nathan Biebel, Christopher Britton, Tim Butzer, John Caravello, Christopher Cloos, Jason Collins, Sherri Conklin, Kevin Craven, Nicole Dular, Kyle Dickey, Mike Garcia, Matthias Jenny, Peter Jones, Kevin Lande, Lauren Leydon-Hardy, Juliana Lima, Corey McGrath, Robert McIntyre, David Mokriski, Honey Nagakura, John Nosan, Charles Perkins, Richard Sandlin, Matthew Sayball, and Daniel Story.

I want to thank all the friends I have made before graduate school. Among them, I especially want to thank Tanya Kostochka. Tanya has been among my very best friends since my 2nd year of undergrad. We have had many adventures, laughs, cries, and amazing meals together since. One day, we'll find a replacement for Asia Wok. Or we'll just visit Asia Wok again and see if it's still the same.

I want to thank my family. Special mention goes to my brother, Dale Andrew Batoon, and to my parents, Ethel Batoon and Arnel Batoon. Without their love and support, I wouldn't have the character to strive for what is right, despite how hard and uncertain that may be. I love them so much and I hope this dissertation does right by them.

Finally, I want to thank my spouse, Chelsea Brown. Her love and companionship has seen me through my highest highs and my lowest lows at UCSB. Her intellectual verve makes me smile and laugh harder than I ever have before. I am so excited to see what the rest of our life together has in store for us.



# Curriculum Vitæ

## Arnel Blake Escobal Batoon

### Education

2024	Ph.D. in Philosophy (Expected), University of California, Santa Barbara.
2016	C.Phil in Philosophy, University of California, Santa Barbara.
2013	M.A. in Philosophy, Brandeis University
2010	B.A. in Philosophy, Brandeis University

### Professional Employment

2023–Present	Lecturer, California Polytechnic State University, San Luis Obispo
2021–Present	Research Mentor, Polygence
2014–2023	Teaching Assistant, University of California, Santa Barbara
2015	Teaching Associate, University of California, Santa Barbara
2011–2013	Teaching Assistant, Brandeis University

## Abstract

Unified Agency, Rational Lies, and the Murderers at the Door

by

Arnel Blake Escobal Batoon

Ambitious presumptivism says that all our testimony based beliefs are on-balance immediately and defeasibly warranted. The rational deception objection says that ambitious presumptivism is not true because it is sometimes rational for a speaker to assert lies rather than truths. One logically possible reply is to argue that it is never rational for a speaker to assert lies rather than truths. In this essay, I develop such a non-conciliatory response to the rational deception objection.

In chapter 1, I explain ambitious presumptivism and the rational deception objection. I identify Kant's prohibition against lying as a historical predecessor to the non-conciliatory response to the rational deception objection. I then identify Burge as the heir apparent to a neo-Kantian non-conciliatory response to the rational deception objection. In chapter 2, I explain my interpretation of how Burge is heir apparent to a neo-Kantian non-conciliatory response. I call Burge's response the "functional unity" argument.

In chapter 3, I defend my attribution to Burge of the functional unity argument. In chapters 4 and 5, I defend the functional unity argument itself from the most influential objections raised against it. In chapter 6, I defend the functional unity argument from the classic murderer-at-the-door objection that dogged Kant's prohibition against lying.

# Contents

<b>Curriculum Vitae</b>	<b>viii</b>
<b>Abstract</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Ambitious Presumptivism Explained . . . . .	1
1.2 The Ideal Rationality of the Occasional Lie . . . . .	13
1.3 Non-conciliatory responses to the rational deception objection . . . . .	23
1.4 Kant: A Historical Precursor . . . . .	25
1.5 Burge: The Contemporary Standard Bearer . . . . .	30
1.6 Roadmaps . . . . .	31
<b>2 The Functional Unity Argument</b>	<b>33</b>
2.1 Parsing the Functional Unity Argument . . . . .	34
2.2 Is The Functional Unity Argument Nonempirical? . . . . .	50
<b>3 Objections and Replies, Round One</b>	<b>56</b>
3.1 Objection One: Rejection-Out-of-Hand . . . . .	56
3.2 Reply to Objection One . . . . .	57
3.3 Objection Two: “Burge Believes RDO1” . . . . .	58
3.4 Reply to Objection Two . . . . .	62
<b>4 Objections and Replies, Round Two</b>	<b>68</b>
4.1 Objection Three: The “Non-Starter” Objection . . . . .	69
4.2 Reply to Objection Three . . . . .	70
4.3 Objection Four: “Equivocating” on the Senses of Rationality . . . . .	75
4.4 Reply to Objection Four . . . . .	79
4.5 Round Two Debrief . . . . .	83
<b>5 Objections and Replies, Round Three</b>	<b>85</b>
5.1 Objection Five: Graham’s Coherent Doubts . . . . .	85
5.2 Reply to Objection Five . . . . .	90

5.3	Rejoinder to the Reply to Objection Five . . . . .	104
5.4	Response to the Rejoinder to the Reply to Objection Five . . . . .	109
5.5	Some Closing Remarks . . . . .	127
<b>6</b>	<b>Burge, Kant, and the Murderers at the Door</b>	<b>130</b>
6.1	The Murderers At The Door . . . . .	131
6.2	Kant and the Murderers at the Door . . . . .	133
6.3	Burge and the Murderers at the Door . . . . .	139
6.4	Against “Instrumental First” Replies . . . . .	141
6.5	Against “Korsgaardian Unity” Replies . . . . .	142
6.6	Against “Goldilocks Critical Reasoning” Replies . . . . .	146
6.7	The “Deviant Contexts” Reply, Part One: Shiffrin on Suspended Contexts	150
6.8	The “Deviant Context” Reply, Part Two: Potential Resistance considered	153
6.9	The “Deviant Context” Reply: A Final Defense . . . . .	158
6.10	Conclusion . . . . .	160
	<b>Bibliography</b>	<b>161</b>

# Chapter 1

## Introduction

Ambitious presumptivism says that all our testimony based beliefs are default defeasibly warranted. The rational deception objection says that ambitious presumptivism is not true because it is sometimes ideally rational for a speaker to assert deceitfully rather than truthfully. In this essay, I critically evaluate a promising response on ambitious presumptivism's behalf.

I start with a primer explaining ambitious presumptivism for the uninitiated. After that, I explain the rationale for thinking that it is sometimes ideally rational for a speaker to assert lies rather than the truth. I then articulate some reasons for why a proponent of ambitious presumptivism should reject the occasional ideal rationality of asserting lies rather than the truth. I introduce Kant's ethics as a historical precedent for this strategy. I then point to Burge's celebrated defense of ambitious presumptivism as an underappreciated contemporary instance for this strategy. I then follow that with a roadmap for the rest of the essay.

### 1.1 Ambitious Presumptivism Explained

Readers who are familiar with ambitious presumptivism can comfortably skip this section. This section exists entirely for those who are unfamiliar with ambitious pre-

sumptivism and desire a step by step introduction to its content. I believe that many contributors have made deep insights into the thesis and the problems it solves. I also believe that these insights are often locked behind the technical trappings of this or that proprietary terminology. If I attempt to faithfully recreate such terminology, I invite the distracting objection that the resulting recreation is not faithful. In addition, if I attempt to faithfully recreate someone else's terminology, I might give the reader the mistaken impression that my aims are purely exegetical. In this primer, I plan to sidestep these distractions by stating the view in my own words and without any attempt to ensure fidelity to someone else's proprietary terminology.

I reproduce the core general thesis thus:

**APG** All of a person's testimonial reception based beliefs are immediately and defeasibly on-balance warranted

I parse APG "from left to right", starting with the occurrence of "all". Following that is "a person's testimonial reception based beliefs". Following that is "immediately and defeasibly on-balance warranted".

### **'All' explained**

The occurrence of "all" in APG indicates that APG is a universal generalization. APG says of a domain that if an element of that domain has the property of being a person's testimonial reception based belief, then it also has the property of being immediately and defeasibly on-balance warranted. Intuitively, APG is not true if it is the case that at least one element of the domain is both a person's testimonial reception based belief and not immediately and defeasibly on-balance warranted. Otherwise, the statement is true.

**‘a person’s testimonial reception based belief’ explained**

In APG, the occurrence of “all” is followed by an occurrence of “a person’s testimonial reception based beliefs”. Here I first explain what I mean by the term “a person’s testimonial reception based beliefs”. Then I address some frequently asked questions some may have about my usage of the term.

A person forms such beliefs when deploying or relying upon the “testimonial reception based belief forming process”. The testimonial reception based belief forming process takes as its starting point our practice of giving testimony. A speaker’s testimony is their action of asserting some content as veridical.<sup>1</sup> This is something speakers often do by specifically performing “sentential” assertions. A speaker performs a sentential assertion by uttering a declarative sentence which would, in the context of utterance, semantically express the content they are presenting as veridical. The content expressed in that context of utterance is the content asserted. Speaker’s sentential assertions can be spoken, signed, or written. If the same content is sententially asserted separate by two different speakers, then those are two separate testimonies. This is because actions are individuated in part by who performs them. If one and the same speaker sententially asserts the one and the same content on two temporally distinct occasions, those are two separate testimonies.

A person receives a speaker’s testimony if and only if the person perceives the speaker’s testimony and they comprehend content asserted in the speaker’s testimony. For example, suppose that Aisha asserts that the store is closed. Raya receives that testimony if two things happen. First, Raya perceives Aisha’s testimony, i.e., through her senses, she has a perceptual representation of the event. Maybe Raya hears Aisha’s

---

<sup>1</sup>On at least one mainstream proposal in the philosophy of action, people’s actions are events. So, as I understand it, a speaker’s testimony is an event. For a classic statement of the view that people’s actions are events, see Davidson (2001a).

speaking the assertion. Maybe Raya sees the ad hoc signage Aisha produced which records and reproduces the Aisha's assertion in writing. Other perceptual modalities are possible, given the right combination of a medium for Aisha's assertion and Raya's endowment of perceptual capacities.<sup>2</sup>

Second, Raya comprehends the content asserted in Aisha's testimony. In whatever way Raya perceives Aisha's testimony, Raya also understands Aisha's message. I leave it to psycholinguists and philosophers of language and psychology to settle the details of what it is to comprehend testimony and speech more generally. What matters for present purposes is that there is a principled causal triggering relationship from a person's perceptions of a speaker's testimony to the person's comprehensions of the content of the speaker's testimony. This triggering relationship makes it so that the person comprehends the purportedly perceived speaker's testimony as their asserting some content.

Now let us turn from a person's testimonial reception to their testimonial reception based belief. Earlier, I said that a person forms such beliefs when they deploy or rely upon the testimonial reception based belief forming process. That testimonial reception based belief forming process consists in a person's advancing from receiving a speaker's testimony to accepting its content. A person accepts some testimony's content, in the relevant sense, when the person believes the content of the testimony (at least in some part) because they received that testimony.

For example, imagine that Raya receives Aisha's testimony about the store closure and then Raya believes that the store is closed. Further, suppose that Raya's receipt of Aisha's testimony plays some "causal explanatory" role in prompting and supporting Raya's belief that the store is closed. In that case, it means that Raya's belief that the

---

<sup>2</sup>There might be complications for cases of written or recorded assertions. Maybe the production of an assertion, i.e., the act of writing the assertion down, and the thing perceived, e.g., the finished writing, are two different things, related as cause and effect. Although I'm not sure that this is correct, I'm fairly certain that this is more a problem about the media, rather than the epistemology, of assertion. For the case of written or recorded assertion, simply adjust the description of the process accordingly.



store is closed is a testimonial reception based belief. Again, I leave the details of the causal explanation to psycholinguists and philosophers of language and psychology. What matters is that we have an intuitive idea of what it is to believe something because of one person's testimony, rather than another person's testimony or some other non-testimonial route.

I describe a person's use of the testimonial reception based belief forming process in terms of the person's transitioning from receiving a speaker's testimony to the acceptance of that testimony's content. I do not mean to imply that the process is fully automatic. People do not necessarily transition from testimonial reception to testimonial reception based belief. Many people accept far fewer testimonial contents upon testimonial receipt than other people. However, when they do accept testimonial contents, those beliefs are testimonial reception based beliefs.

### **'Immediately and Defeasible On-Balance Warranted Belief'**

Here again is ambitious testimonial presumptivism:

**APG** All of a person's testimonial reception based beliefs are immediately and defeasibly on-balance warranted.

I have already elaborated on the occurrences of 'all' and 'a person's testimonial reception based beliefs'. Now, I turn to the last piece, the occurrence of the phrase 'immediately and defeasibly on-balance warranted'. I first talk about what it is for a belief to be warranted, then I move through the distinctions between immediately versus mediately warranted belief, defeasibly versus indefeasibly warranted belief, and on-balance versus *pro tanto* warranted belief. Along the way, I also explain the difference between general and particular warrant, the distinction between a person's body of warrant and their stock of evidence, and the relationship between a person's warranted beliefs and their

stock of knowledge.

A person's warranted beliefs are those of their beliefs that are ideally rational as proper attempts at thinking truthfully. Such beliefs are so rational because they are proper attempts at thinking truthfully. For example, if Raya forms a belief about what is going on outside her apartment by flipping a coin, the resulting belief is not a warranted belief. Raya's coin flip based belief about events outside her apartment is not a warranted belief because it is not ideally rational as a proper attempt at thinking truthfully. This verdict seems right because a coin flip based belief about particulars in the immediate environment is not a proper attempt at thinking truthfully. Such coin flip based beliefs are not, in general, proper attempts at thinking truthfully, because there is not a principled explanatory relationship between coin flip based beliefs about particulars in the immediate environment and their subject matters.

For contrast, suppose that instead of coin flips, Raya deploys her sensory perceptual capacities to settle her questions about what's going on just outside her apartment. As a result, she would form perception based beliefs about the events outside her apartment rather than coin-flip based beliefs. Those perception based beliefs are warranted because they are proper attempts at thinking truthfully. This makes sense because there's a principled explanatory relationship between the types of particulars present in the types of immediate environment she might find herself in, the types of sensory perceptions her perceptual systems would generate about those types of particulars, and the types of beliefs she bases off the contents of sensory perceptions. More generally, there is a principled explanatory relationship between the types of beliefs that a person may form using the perception based belief forming process and the (types of) subject matters of those beliefs. It is in virtue of this principled explanatory relationship that Raya's belief about events outside her apartment are ideally rational as proper attempts at thinking truthfully.

Here is how talk of a person's warranted belief relates to talk of whether they have a body of warrant for that belief. A person's warranted beliefs, as proper attempts at thinking truthfully, are the person's applications of principled explanatory relationships between beliefs and their subject matters. In this locution we can identify both a general component and a particular component.

First the general component. The principled explanatory relationships between the types of beliefs a person may form and the subject matters of those beliefs comprise a their *general warrants* to use the belief forming methods associated with those principled relationships. General warrants, in virtue of being constituted by principled explanatory relationships between certain types of beliefs and their subject matters, are not particular to any specific person's situation. Any person who has the associated belief forming process in their repertoire is a person who enjoys the general warrant to the beliefs output by that process. Sometimes, when there's a need to distinguish those belief forming processes constituted by principled explanatory relationships between certain types of beliefs and their subject matters, and those that are not, I will refer to the latter belief forming processes as "channels for warrant".

Second, the particular component. The particular component is comprised of a person's particular warrants. A person's particular warrants are composed of specific applications of their general warrants to their particular contexts. You might, for instance, enjoy a general warrant to a certain sort of belief but not be in a position where using it would get you an instance of the relevant belief. For example, Raya might have a general warrant to her perception based beliefs, but lack particular warrant for a perception based belief about what's happening outside. She might lack such a warrant because, given her spatio-temporal positioning, she's not in a position to bring her sensory perceptions to bear on what is happening outside her home.

Alternatively, Raya might be in a situation to enjoy both a general warrant to her

perception based beliefs and particular warrant for such perception based beliefs regarding what is happening outside her home. She might do so in a situation because in that situation she not only possesses an instance of the perception based belief forming process, she is also in a spatio-temporal position to bring that capacity to bear on particulars outside her home. She might, for instance, be standing just outside her home.

The two examples above illustrate that for a person to have warrant for a belief, specific elements of a person's total body of warrant must align. Specifically, they need to have a general warrant to form a certain type of belief, and they need to have particular warrant, in the form of situational conditions, to betoken that type of belief in that scenario. Accordingly, a person's total body of warrant is comprised of their general warrants together with the particular ways those warrants combine with their situation.<sup>3</sup>

My way of talking about a person's warrants and a person's warranted beliefs allows us to distinguish between a person's body of warrant for their beliefs and their stock of evidence. A person's body of warrant is the set of properties and relationships that determines which of a person's beliefs are proper attempts at thinking truth. This is traditionally called a person's propositional warrant because it is determines the set of propositions the person would be warranted in believing, should they form those beliefs in the right way. A person's stock of evidence consists of which warranted beliefs they actually form. Put plainly, a person's stock of evidence just is their stock of warranted beliefs. A person might have warrant for a belief, but not deploy the warrant. Alternatively a person might form a belief they have warrant for, but not in a way that deploys that warrant.<sup>4</sup> For example, Raya may have the evidence required to deductively infer

---

<sup>3</sup>I borrow the phrase "body of warrant" from Burge (1993). Burge further distinguishes between a person's "proprietary" and "extended" bodies of warrant. This is a distinction meant to capture how some elements of a person's body of warrant are "local" to them while others are "imported" in through the reception of testimony. This is a deeply vexing issue, but it's not the issue I'm examining in this essay.

<sup>4</sup>I think this distinction echoes Pryor's distinction between having justification for a proposition and appropriately having a belief in that proposition. See Pryor (2005).

something and yet not perform the deduction. Instead, she might flip a coin and form her beliefs based on the coin flip. Her resulting belief would be unwarranted because she based it on a coinflip rather than her warrant. As such, it wouldn't belong among her total evidence.<sup>5</sup>

My identification of a person's actual warranted beliefs with their stock of evidence is deliberate. When we think about what it is for a person's beliefs to be genuine evidence, we think of those as having the potential to actually rationally support the acceptance of some further claim. In short, a person's beliefs are evidence when they genuinely have the potential to support further beliefs as proper attempts at thinking truthfully.

I characterized a person's warranted beliefs as those that were ideally rational as their proper attempts at thinking truthfully. It is possible for some of the person's proper attempts at thinking truthfully, i.e., their warranted beliefs, to be successful as proper attempts at thinking truthfully. Such warranted beliefs are successful attempts

---

<sup>5</sup>In embracing a distinction between a person's warrant and their evidence, I oppose coherentism and evidentialism about warrant. Coherentism about warrant is the view that a person's beliefs are warranted to the degree to which their beliefs are interconnected by a web of introspectively accessible explanatory connections. Important coherentist statements include Bonjour (1985), Davidson (2001b), Fricker (1994), and Sellars (1956). Evidentialism about warrant is the view that a person's beliefs are warranted to the degree to which their beliefs are supported by their evidence. Important evidentialist statements include Conee and Feldman (2004).

On both the coherentist and evidentialist views, a person cannot be wrong about what their evidence is. On both views, a person's warrant is reducible to their evidence. Both views mainly disagree on what evidence is. The coherentist thinks that a person's evidence consists entirely in their warranted beliefs. The evidentialist thinks that a person's evidence at a time consists entirely in which of their mental states and events are introspectively accessible at that time.

Against both views, I think that a person can be mistaken about their evidence. I think this because I think a person can be mistaken about their warrant. I do not think that a person's warrant must consist in mental states that are introspectively accessible to them. The principled explanatory relationship between a person's perceptions of particulars, their perception based beliefs about those particulars, and the particulars themselves is not necessarily the sort of thing that is introspectively accessible to a person at a time. A child can have a warranted perception based belief that their cat is black and white without being in a position to veridically cognize the general causal relationships that make their perception based belief a proper attempt at thinking truthfully. Indeed, they can have such a warranted belief while being entirely ignorant or mistaken about how perception based beliefs are proper attempts at thinking truthfully. For such people, it is the explanatory relationship between their perceptions, their perception based beliefs, and those belief's subject matter, not their successful exercises of introspection, that constitute their warrant for their perception based beliefs. For elaborations of this idea, see Burge (2003) and Graham (2012). These proposals develop on ideas suggested by Goldman (1979).

at thinking truthfully because they are in fact true. At least some of these warranted true beliefs exhausts the person's stock of knowledge.

With that broad gloss on warrant and warranted beliefs out of the way, I turn to the distinction between immediately and mediately warranted belief. The distinction between immediately and mediately warranted belief is a distinction between the ways that a belief's warrant is related to their total body of evidence. A person's belief is immediately warranted to the extent that some of the person's warrant for that belief is not inferentially derived from their total evidence. A person's belief is mediately warranted to the extent that some of the person's warrant for the belief is inferentially derived from their total evidence. In the latter case, it is not uncommon to talk of the person's total evidence providing support or reason for the belief.

Consider, for example, two ways that Raya's belief that the tree leaves outside her apartment are red. She might believe that the tree leaves outside her apartment are red because she went out and looked at them. In this case, her warrant consists in her use of her perception based belief forming process. The explanatory connection associated with a person's perception based beliefs is one that exists regardless of a person's total evidence. Intuitively, this would mean that, insofar as Raya's belief about the leaves is a perception based belief, it is an immediately warranted one.

Alternatively, Raya might believe that the tree leaves outside her apartment are red because she inferred it from her evidence. Suppose her evidence includes her knowledge that those trees are ones whose leaves change color in autumn and her knowledge that it is autumn. Suppose she uses her knowledge to infer, and hence believe, that the tree leaves outside her apartment. In that case, her warrant for that belief involves her making an inference from her evidence. If so, her inference based belief is mediately warranted.

What these examples illustrate is that these modes of warranted belief, immediacy and mediacy, are tied to the type of belief forming processes used to support the belief.

Suppose that Raya first performs the inference described above and then follows that up with perceptual observation. In that case, her belief about the tree leaves *both* mediately and immediately warranted. As an evidential inference based belief, it is mediately warranted. As a perception based belief, it is immediately warranted. Her warrant is overdetermined.<sup>6</sup> Raya's overdetermined warrant consists in at least two independent *channels* of warrant. A channel of warrant is a belief forming process (type) that is associated with a general warrant, i.e., a principled explanatory connection between inputs, output beliefs, and the subject matter of those beliefs. Perception is one channel. Evidence based inference is another channel.

Now let's turn from the distinction between immediately and mediately warranted belief to the distinction between defeasibly and indefeasibly warranted belief. Whereas the distinction between immediately and mediately warranted belief is about whether a person's belief draws inferential warrant from their total evidence, the distinction between defeasibly and indefeasibly warranted belief is about whether a person's belief suffers defeat from their total evidence. A person's warranted belief is defeasibly warranted if and only if their warrant for their belief can be either undermined or overridden by their total evidence. Otherwise, the person's warranted belief is indefeasibly warranted.

For example, Raya might have a warranted perception based belief that the tree leaves outside her apartment are red. That belief would be defeasibly warranted because her body of evidence could undermine that belief. In particular, she might already have, or later acquire, evidence to the effect that a crucial aspect of her perceptual based belief forming process is compromised.

Examples of indefeasibly warranted beliefs are harder to come by. Perhaps Raya believes a seemingly trivial analyticity, like the proposition that bachelors and bachelorettes are unmarried people. Assuming that Raya genuinely understands the concepts deployed

---

<sup>6</sup>For a brief statement of the "reality" or "validity" of overdetermined warrant, see Burge (2013).

in that proposition, it seems correct to say that there is no possible total body of evidence that could marshal against the truth of that claim. The truth of such a proposition depends instead on the meanings of its component concepts.

I now turn from the distinction between defeasibly and indefeasibly warranted belief to the distinction between on-balance and merely *pro tanto* warranted beliefs. A person's belief is on-balance warranted when their warrant for the belief is adequate to ideally rationalize the belief as a proper attempt at thinking truthfully. A person's belief is *merely pro tanto* warranted when the person's warrant for the belief is not adequate. For example, if Raya uses axioms and basic definitions to form a deduction based belief in a mathematical theorem and her beliefs in those axioms and basic definitions is warranted, then her deduction based belief in that mathematical theorem is on-balance warranted. Alternatively, if Raya has not yet performed the deduction, but has tutored reliable intuitions that she could perform such a deduction given enough material support, she has a merely *pro tanto* warranted belief.<sup>7</sup>

Having guided us through warrant, immediacy, defeasibility, and on-balance, I can fully elaborate upon what I think APG means. Once again, here is APG, the statement of ambitious testimonial presumptivism:

**APG** All of a person's testimonial reception based beliefs are immediately and defeasibly on balance warranted.

APG essentially says that for each element in the domain (e.g., the actual world), if that element is a person's testimonial reception based belief, then it has the property of being adequately ideally rationalized as the person's attempt at thinking truthfully in a way that doesn't draw inferential support from the person's evidence (immediacy) and in a

---

<sup>7</sup>With these examples, I am helping myself to a common sense view that we don't have on-balance warrant for theorems of arithmetic without having deduced them as theorems from propositions we have on-balance warrant to believe.



way that can be defeated by the remainder of the person's total evidence. Put another way: A person's merely receiving a speaker's testimony is enough for the person's belief to be defeasibly warranted.

As stated, APG describes a general warrant. It is the kind of warrant that is available to anyone with the capacity to form beliefs by receiving a speaker's testimony. It comprises everyone's warrant to rely on testimony as a source of warranted beliefs. The general warrant described gets applied to a person's individual beliefs when, in the context of belief formation and belief maintenance, it satisfies the condition of being a testimonial reception based belief. APG, together with those situational conditions, constitutes an individual person's specific warrant to specific testimonial reception based beliefs.

## 1.2 The Ideal Rationality of the Occasional Lie

Having worked through testimonial presumptivism, as encapsulated by APG, I now turn to the rational deception objection. Put in the broadest possible terms, the rational deception objection is that APG is not true because it is on some occasions on-balance ideally rational for a speaker to assert lies rather than truths. Here's my initial reconstruction of the basic argumentative structure of the rational deception:

**RDO1** It is on some occasions on-balance ideally rational for a speaker to  
assert lies rather than truths

**RDO2** If it is on some occasions on-balance ideally rational for a speaker to  
assert a lie rather than the truth, then APG is not true

**RDO3** So, APG is not true

Here is what RDO1 says. It tells us that there is at least one occasion in which it is

on-balance ideally rational for a speaker to deceive. RDO2 says that the fact described in RDO1 undermines the truth of APG. The two of these combine to form a deductively valid argument for RDO3. Accordingly, the best objection to the rational deception objection would cast doubt on either RDO1 or RDO2.

RDO1 and RDO2 feature quantification over occasions. At an intuitive level, words like ‘circumstance’, ‘context’, ‘environment’, ‘event’, ‘occasion’, ‘outcome’, ‘scenario’, and ‘situation’ refer to something like some part of some possible world for some interval of time. For my purposes, an interval of time can be as short as the smallest instance of time or as long as all eternity. Similarly for my purposes, a part of a possible world can be as small as whatever the correct mereology of possible worlds allows, and it can be as large as a whole possible world itself. Hence, I will use the terms ‘circumstance’, ‘context’, ‘environment’, ‘event’, ‘occasion’, ‘outcome’, ‘scenario’, and ‘situation’ to refer to parts of worlds at intervals of times, where the sizes of those world-parts and time-intervals are left fully variable. Accordingly, I will use the terms ‘circumstance’, ‘context’, ‘environment’, ‘event’, ‘occasion’, ‘outcome’, ‘scenario’, and ‘situation’ interchangeably.<sup>8</sup>

For reasons I make clear in the next section, I focus entirely on an objection to RDO1 in this essay. To that end, I want to give a sustained explanation for why RDO1 seems initially plausible.

---

<sup>8</sup>I acknowledge that there are significant debates over the metaphysics of possible worlds, the metaphysics of time, the metaphysics of events, decision theory, event semantics, possible worlds semantics, and situation semantics. These philosophical debates have imbued words like ‘circumstance’, ‘context’, ‘environment’, ‘event’, ‘occasion’, ‘outcome’, ‘scenario’, and ‘situation’ with technical proprietary meanings and presumptions. Some positions in some of these debates constitute significant advances in philosophical and scientific understanding. Some of them do not. I do not want to take a position on any of them in either way. Do not misunderstand me, I do not presume to think that a complete philosophical or semantic theory can be truly neutral on these debates. Perhaps the way I philosophize might push me towards one set of stances on those debates versus others. Nevertheless, none of those debates are debates about the epistemology of testimony. I’m happy to keep those debates separate from debates about the epistemology of testimony. Just as a biologist may study evolution without having an opinion on how their theory is derivable, if at all, from first principles in quantum mechanics, I think an epistemologist may study the epistemology of testimony without having endorsed a fully worked out formal semantics backed up by a metaphysics of worlds, times, and events.

I think the best recipe for seeing why someone might accept RDO1 is to proceed by examples. To do this, you need to do two things. First, you need to offer a plausible account of what it is for an action to be on-balance ideally rational. Second, you need to provide an example of a case where according to that account, it is on-balance ideally rational for a person to assert a lie rather than a truth.

One popular account of on-balance ideally rational is a subjective desire satisfaction account. On this view, a person's action is on-balance ideally rational if and only if they believe that their action would satisfy their on-balance desires. For example, suppose that Aisha's on-balance desire is to have spinach for dinner. Suppose also that she is in a position to perform the action wherein she goes to the store to buy spinach. Let's also suppose that the Aisha believes that such an action would satisfy her desire to have spinach for dinner. According to the subjective desire satisfaction account, it is on that occasion on-balance ideally rational for Raya to perform that action.

Armed with a subjective desire satisfaction account of on-balance ideally rational action, I think that we can easily come up with a case where a speaker's asserting a lie rather than a truth would promote the satisfaction of their desires. For example, we can imagine that Aisha believes that the prospective labor contract negotiated by her pro-business union is not adequate for her constituents to live on. Suppose that it's Aisha's on-balance desire to have a lucrative career at the union's international offices. Suppose Aisha believes that if she were to tell her rank and file union constituents that the contract is adequate for them to live on, she would be lying to them. Finally, suppose that Aisha believes telling such a lie would satisfy her on-balance desires to have a lucrative career at the union's international offices. According to the subjective desire satisfaction account, it would be on-balance ideally rational for Aisha to tell the lie rather than the truth.

Here is another example. Suppose that Aisha has an on-balance desire to be the only applicant among her class for a certain job. Suppose that Aisha believes she is in a

position to successfully lie to her classmates about who is eligible to apply. And suppose that Aisha believes that such a lie would be enough to satisfy her desire to be the only job applicant. According to a subjective desire satisfaction account, it is on that occasion on-balance ideally rational for the Aisha to assert a lie rather than the truth.

I have specified at least two logically possible scenarios which, together with the desire satisfaction account, entail that on some occasions it is on-balance ideally rational for a speaker to assert a lie rather than a truth. These logically possible scenarios might not cite any specific examples, but I think it's going to be hard to deny that there are plenty of real life scenarios like these ones. Of course, you might wonder whether the account of on-balance ideally rational action that I have offered is sophisticated enough to handle the intricacies of real life. For those who worry about that, I offer another popular account of on-balance ideally rational action.

Another popular account of on-balance ideally rational action is subjective decision theory.<sup>9</sup> Subjective decision theory ties what's rational for a person to do on some occasion with what expected utilities they "associate" with that action. The core idea is that what's rational for a person to do is the action with the greatest expected utility. Let's breakdown that idea starting with the notion of utility, and then building towards the notion of expected utility.

First, the utility a person associates with an action represents the general pay-off or reward associated with that action. The person's utility for an action is calculated by summing together the particular pay-offs associated with particular pairings of that action with different possible outcomes.

Suppose Aisha and Bianca are playing a simple card game. The game is that Aisha

---

<sup>9</sup>Classic statements of subjective decision theory include Savage (1954) and Jeffrey (1965). These proposals are subjectivist in that they only propose standards for the internal structure of a person's body of beliefs. They do not impose further for how a person's beliefs relate to the subject matter of those beliefs.

draws a card from a standard 52 card deck and Bianca has to call out which color the card's suit is before Aisha reveals it. Because it is a standard 52 card deck, the cards only come in two colors, black and red. There are 26 black cards and 26 red cards. If Bianca calls out the correct color, she gets 3 dollars. If Bianca calls out the wrong color, she loses 3 dollars. According to this set up, the utility of calling black is 0 dollars. That's because the pay-off for calling black, and the card being black, is 3 dollars while the pay-off for calling black, and the card being red, is -3 dollars. Summing those pay-offs together, the result is dollars.

A person's expected utility for an action represents the general pay-off associated with that action, as modulated by their point of view of how probable the different outcomes are to occur. A person's expected utility is determined by summing together the expected pay-offs of pairings of outcomes to the action in question. The expected pay-off of a particular action-outcome pairing is determined by multiplying the particular pay-off of that pairing with the person's probability or degree of confidence in the outcome occurring.

Suppose that since the start of the game, Aisha has drawn 13 cards, all 13 of the cards drawn have been black, and that Bianca knows of all that. Accordingly, from Bianca's point of view the probability she assigns to the next card being red is  $2/3$ . The probability she assigns to the card being black is  $1/3$ . As a consequence, Bianca's utility for calling black on the next draw is not the same as her expected utility for calling black on the next draw. Like we noticed before, for Bianca, the utility of calling black is 0 dollars. However, for Bianca, the expected utility of calling black would be -1 dollar. In contrast, for Bianca the expected utility of calling red on the next draw is 1 dollar.

Subjective decision theory espouses a sufficiency claim. It goes like this:

**SDT** On all occasions, if a person's action has the greatest expected utility for them,

then their action is ideally on-balance rational.

SDT seems plausible insofar as it is supported by the intuition right to say that what's rational for you to do is determined by what you think best benefits you. SDT, supported by the apparatus of subjective decision theory, gives a plausible reconstruction or model of that intuition. This theory addresses a worry about how people can rationally proceed when they have incomplete evidence about how the world is. A plausible insight is that person's thoughts or feelings about which outcomes are more probable constitute their attempts at navigating such uncertainty. For all intents and purposes, SDT uses that insight to amend subjective decision theory to address the worry regarding how people have indecisive evidence about the world.

An example may help illustrate SDT. Suppose that the Aisha and Bianca's game has progressed through thirteen draws as described in the previous paragraph. That is, thirteen of the fifty two cards have been drawn and all of those drawn cards have been black. The remaining information has been amalgamated into the table labelled "Bianca's Next Draw". Accordingly, the expected utility of calling black on the fourteenth draw

Action	Next is Black (1/3)	Next is Red (2/3)	Expected Utility
Call Black	3	-3	-1
Call Red	-3	3	1

Table 1.1: Bianca's Next Draw

is -1 dollar, and the expected utility of calling red on the fourteenth draw is 1 dollar. According to SDT, because calling red on the fourteenth draw bears greater expected utility for Bianca than calling black on the fourteenth draw, it is on-balance ideally rational for Bianca to perform the former action over the latter.

Armed with subjective decision theory and SDT, it's not hard to see how it is sometimes on-balance ideally rational for a person to assert lies rather than truths. We can quickly conjure up a possible occasion on which it is on-balance ideally rational for a

person to assert lies rather than truths. First, let us suppose that it turns out that the pharmaceutical megacorporations Fitzer and BjornSkill worked together to create a vaccine for the Killer Cold. Further, let us suppose that their vaccine is scientifically known to be a safe and effective enough to prevent hospitalization from the Killer Cold.

In addition, let us suppose that Carla Tuckersdottir has learned about Fitzer-Bjornskill vaccine for the Killer Cold. In talking to people, if she says anything about Fitzer-Bjornskill vaccine at all, she has two options. First, she can assert the truth about the vaccine, namely that it is a safe and effective enough vaccine for Killer Cold. Second, she can assert a lie about the vaccine, either that the vaccine is not safe enough or is not effective enough.

In addition, let us suppose that for Carla Tuckersdottir, the potential outcomes as she understands them are as follows. The first is that her audience is more receptive to the truth rather than a lie. The second is that Carla's audience is more receptive to a lie rather than the truth.<sup>10</sup> By Carla's lights, there are no other possible outcomes.

Further, let us suppose that from Carla's point of view, the outcome that her audience is more receptive of her asserting a lie rather than the truth is twice as probable as the outcome that her audience is most receptive to her asserting the truth rather than a lie. Let's also grant that Carla's outlook is grounded in her knowledgeable of what sorts of messages her audience is most receptive to. That is, she thinks that her audience will most likely be receptive to messages that support an anti-vaccination position.<sup>11</sup>

Further, let us suppose that Carla sees her utility in terms of net changes to the number of her audience, i.e., in the number of people who continue to talk to her or

---

<sup>10</sup>These are Carla's descriptions of the outcomes. It is not that the audience wants to be lied to versus told nothing or told the truth. We can charitably assume that the audience wants to be told the truth. However, by Carla's estimation, the contents of certain types of messages rather than others will "jive" or "vibe" better with the audience.

<sup>11</sup>Such a situation, where an audience is most receptive to an anti-vaccination position, is not merely logically possible. In the real world, there are many communities where people harbor or even endorse an anti-vaccination position.

stand on the look-out for what she has to say. For Carla, an increase in that number is a positive pay-off. For Carla, a decrease in that number is a negative pay-off. Finally, a no change in that number either way is a neutral pay-off. This seems a plausible enough measure utility insofar it is intuitive that speakers want to be engaged with and listened to rather than ignored.

Now let's make suppositions about how Carla understands the pay-offs for each of the outcomes relative to her available actions. For the outcome of her audience being more receptive towards the truth than a lie, relative to the action of her asserting the truth rather than a lie, the pay-off is a net increase her audience by 3. For the same outcome relative to her action of asserting a lie, the pay-off is a net decrease in her audience by 3. For the outcome of her audience being more receptive to a lie than the truth, relative to the action of her asserting the truth rather than a lie, the pay-off is a net decrease in her audience by 3. For the same outcome, but relative to her action of asserting a lie, the pay-off is a net decrease in her audience by 3.

All of the foregoing stipulations have been amalgamated into a table. I have labelled it "Carla's Point of View". This table presents how Carla's situation is one where Carla's

Action	People like Truth (1/3)	People like Lie (2/3)	EU
Assert Truth	3	-3	-1
Assert Lie	-3	3	1

Table 1.2: Carla's Point of View

asserting a lie about the Fitzer-Bjornskill vaccine is the action with the greatest expected utility. According to **SDT**, that suffices for it to be on-balance ideally rational for Carla to assert a lie about the Fitzer-Bjornskill vaccine rather than the truth. Existentially generalizing from this case, we get RDO1, i.e., there is at least one occasion where it is on-balance ideally rational for a speaker to assert a lie rather than the truth.

I have discussed two accounts of on-balance ideally rational action: the subjective



desire satisfaction account and subjective decision theory. Both accounts are influential examples of such accounts. Further, I have discussed how on each account there is a situation where it is on-balance ideally rational for a speaker to assert a lie rather than the truth. Although there are important differences regarding how these individual accounts reach those verdicts, what matters is that there are ways of further articulating whatever intuition supports RDO1. That intuition is that what is on-balance ideally rational for a person to do is what, in some specific sense, benefits them or furthers their aim; in that light, RDO1 is initially plausible.

A reader might object that my explanations of RDO1's plausibility relies on overly subjective accounts of what makes an action on-balance ideally rational. Perhaps they feel as though what's on-balance ideally rational for a person to do isn't so much a matter of how a person views their actions bring about their ends so much as whether their actions actually bring about their ends. Recall the earlier example pertaining to Aisha's on-balance desire to have spinach for dinner. In that example, Aisha' believed that her action of going to the store would satisfy that on-balance desire. The reader would say that what matters is not what Aisha believes of that action, but instead what that action would actually accomplish.

I think that this idea can be easily accommodated. We can straightforwardly envision an objective desire satisfaction theory according to which a person's action is on-balance ideally rational if and only if that person knows their action would satisfy their desires. This is would count as an objective account of on-balance ideally rational action because knowledge is objective. When a person knows something, it's not just that they believe it obtains. What that person believes is actually true and the resulting belief of theirs stands in a warrant supporting principled explanatory relationship to it.

Armed with an objective desire satisfaction account, I think I can simply amend the original examples so that the person knows that the action in question would satisfy their

on-balance desire. For example, I can amend the example of Aisha's on-balance desire to have spinach for dinner so that she knows that her action of going to the store would satisfy her on-balance desire to have spinach for dinner.

Similarly, let's amend the example where Aisha's on-balance desire is to have a career at the pro-business union's international office. Amend it so that she knows that advocating for the pro-business contract would constitute lying and amend it so that she knows that lying would satisfy her on-balance desire to have a career at the union's international offices. So amended, it follows from the objective desire satisfaction account that this is an occasion in which it is on-balance ideally rational for a person to assert a lie rather than the truth.

A similar way to amend subjective decision theory is suggested by Williamson (2000). On his view, a person's beliefs are made rational if they're in accord with what the person knows. He extends this to decision theory by focusing on "evidential probabilities". A person's evidential probabilities are their point of views towards which outcomes are likely, but tempered by what they know. Let's call the resulting decision theory 'Evidential Decision Theory'.

I think it's trivial to generate an example where Evidential Decision Theory would return the verdict that it's on-balance ideally rational for a person to assert a lie rather than the truth. Take the example of Carla Tuckersdottir, but amend it so that in light of what she knows, the outcome wherein her audience is more receptive to a lie rather than the truth is twice as probable as the outcome whereing her audience is more receptive to the truth rather than a lie. Everything else stays the same. In that scenario, Evidential Decision Theory would yield the veridct that it's on-balance ideally rational for a person to assert a lie rather than the truth.

You might think that when we say that what's rational for a person to do is objective rather than subjective, we mean that what's rational for a person to do is not solely

determined by what they believe about their actions or what they want. On such a view, certain actions are rational (or irrational) regardless of the person's goals or desires. On one way to develop such a view, it is on no occasion ideally on-balance rational for a speaker to assert a lie rather than the truth because there is something inherently defective about the reasons behind those actions. I am sympathetic to such a view. The rest of this essay is dedicated to developing a view that does just that.

### 1.3 Non-conciliatory responses to the rational deception objection

I categorize ambitious presumptivist responses to the rational deception objection based on whether they reconcile ambitious presumptivism with how it's sometimes rational for speakers to assert lies rather than truths. Consider again my broadest formulation of the rational deception objection:

**RDO1** On some occasions, it is on balance ideally rational for a speaker to assert deceitfully rather than truthfully

**RDO2** If it is on some occasions on balance ideally rational for a speaker to assert deceitfully rather than truthfully, then ambitious presumptivism is not true

**RDO3** So, ambitious presumptivism is not true

If you pursue *the conciliatory response*, your goal is to argue against RDO2. If you pursue *the non-conciliatory response*, your goal is to argue against RDO1.

If you are sympathetic to the conciliatory response, you may be sympathetic to a certain picture of rationality. On this picture of rationality, rational agency is about

agents striving to promote their intended goals in light of their available evidence. This is the picture of rationality seems to most clearly support RDO1. If you are sympathetic to both this picture of rationality and ambitious presumptivism, then you must find some way to reconcile the truth of both theses. That is, you must find some way for ambitious presumptivism and RDO1 to both be true together. That’s why I dub it ‘the conciliatory response’. Although I’m broadly sympathetic to it, I want to set the conciliatory response aside for now and develop it in another chapter.

If you are sympathetic to the non-conciliatory response, you might be sympathetic to a certain picture of lying. On this picture, a person’s lying is essentially correlated with a specific kind of feature. This feature, in turn, is essentially incompatible with being rational. Hence, on this picture, it is not true that it’s sometimes rational for a speaker to assert lies rather than truths, because lying is essentially incompatible with acting rationally. If you go this route, then you do not reconcile the truth of ambitious presumptivism with RDO1 because you accept that the truth of one of these theses precludes the truth of the other one. That’s why I dub it ‘the non-conciliatory response’.

I explore the prospects for a non-conciliatory response. I find such a response appealing for its dialectical simplicity. An interlocutor might justify RDO2 in all sorts of ways. If you wanted to give a decisive and exhaustive rebuttal of RDO2, you would have to find decisive rebuttals to each of those different ways that your interlocutor might justify RDO2. After all, RDO2 is only as plausible as the justification you give for it. There are, to be sure, more specific attempts at justifying RDO2.<sup>12</sup> If you advocate a

---

<sup>12</sup>Influential attempts at justifying RDO2 include Fricker (1994) and Faulkner (2011). Simion and Kelp (2020) attempt to exhaustively characterizing the possible ways of attempting to RDO2. They dub the justifications of RDO2 “Source Problems”. Simion and Kelp distinguish two variations on the source problem: “Subjective” Source Problems and “Objective” Source Problems. Subjective Source Problems support RDO2 by bringing to bear considerations linking a person’s warrant for their testimonial reception based beliefs with the person’s having on-balance (critical) reasons for believing the source of the relevant testimony to be reliable. Objective Source Problems support RDO2 by impugning the actual world, *de facto* connection of testimonial reception based beliefs to their subject matter. To these source problems, I would add another possible justification for RDO2 that I would call “mismatch” problems.

non-conciliatory response, you can bypass those justifications. If you advocate a non-conciliatory response, you can ambitiously cut to the heart of the rational deception objection for a decisive rebuttal.

## 1.4 Kant: A Historical Precursor

Kant’s moral philosophy is a prominent historical precursor to the non-conciliatory response.<sup>13</sup> In presenting Kant’s moral philosophy as a prominent pre-cursor to the non-conciliatory response, I am not attributing to him the project of specifically responding to the rational deception objection. After all, the project of responding to the rational deception objection requires concepts and distinctions that entered mainstream discourse after Kant was writing. Further, I have no reason to think that Kant had any specific inklings of these ideas in his writing. I’m simply pointing out that Kant’s work, read with a modern lens, seems to exhibit certain commitments that are indicative of a non-conciliatory response to the rational deception objection.

Let’s talk about how Kant’s moral philosophy strikes me as a prominent historical precursor to the non-conciliatory response. Kant’s approach counts as such because Kant explains moral propriety in terms of rational propriety. Kant does this when he says that nothing in itself is intrinsically moral good other than a good will.<sup>14</sup> When he makes this claim, he is introducing the following package claims. First, the moral worth of an action is dependent on the moral worth of the “will” that produced that

---

My idea is that the truth of RDO1 undermines the principled explanatory relationships that would establish in principle, immediate and defeasible warrant constituting relationships between testimonial reception based beliefs and their subject matters. Accordingly even if a proponent of ambitious pre-emptivism could appeal to a view like psychological anti-individualism to dispute the Source Problem justifications of RDO2, psychological anti-individualism could not be deployed as a defense against the mismatch problem. I explore these issues in a companion piece the present essay.

<sup>13</sup>My sketch of Kant’s stance on lying is informed by chapters 5 and 12 of Korsgaard’s (1996). I take full responsibility for any misreadings and misattributions.

<sup>14</sup>Kant (1785/2018), G 4:393-394

action. Second, the moral worth of the “will” depends on whether that “will” satisfies ideal rational standards. Third, those ideal correct rational standards are formulations of the categorical imperative.

Reconstructing Kant’s moral predecessor of the non-conciliatory response, we get the following chain of reasoning. It is not moral to lie to others because the intention to lie is essentially the result of an ill will. The intention to lie is essentially the result of an ill will because an ill will is an irrational will. An ill will is an irrational will because ill wills endorse intentions that do not pass the rational standards.

From this chain of reasoning, we can extrapolate a rationale against RDO1. For Kant, the intentions to lie essentially issue from an ill will and an ill will is essentially a will that produces intentions that fall short of the rational standards. Further, let’s help ourselves to the claim that occasions where a person’s “willing” fall shorts of rational standards are occasions where it’s not rational for the person to act on those “willings”.<sup>15</sup> Since a person’s intending to lie issues from them willing in conflict with rational standards, it follows that it’s not true that it’s sometimes rational to assert lies rather than truths. For such actions will always, for Kant, come from a will that has failed the standards of rationality.

Clearly, the crucial aspect of Kant’s treatment of the rationality of lying is guided by the idea that lying, at least out of self interest, does not pass the categorical imperative. Recall that there were three formulations of the categorical imperative; the universalizability formulation, the humanity formulation, and the legislator in a kingdom of ends formulation. The categorical imperative, as Kant conceived of it, was a norm on rational “willing” or “intending”. A person whose settled intentions do not pass the categorical

---

<sup>15</sup>I think this wouldn’t go amiss because it seems plausible to me that you explain what’s rational in terms of what satisfies rational standards. On Kant’s view, ill wills are wills that allow for plans and actions that don’t satisfy rational standards. To me, this means that if it’s not rational for a person to do something, then the action comes from an ill will.

imperative is a person who falls short of being rational.

According to the universalizability formulation, the moral worthiness of an (intended) action is determined by whether a person could “will it” as a universal law.<sup>16</sup> The courses of actions that could not be intended as universal laws are the ones that would lead to significant break downs if intended as universal laws. For example, consider the moral worthiness of a shopkeeper not giving fair change during a business transaction. The universalizability test asks us to consider whether this type of action could be intended as a universal law without breaking down. Intuitively, this type of action could not. If everyone did this as a matter of universal law, the associated institution of business exchanges would fall apart. Hence, the Shopkeeper’s action of not giving fair change during business transactions is not the kind of action that is morally worthy.

Similarly, consider the action of a speaker’s asserting a lie rather than the truth for self-serving reasons. Kant thinks that this too could not be willed or intended as a universal law. Allegedly, asserting a lie rather than the truth could not be willed as a universal law because it would allegedly lead to a breakdown in the practice of asserting. And insofar as the institution of testimony is built on the practice of asserting, it would seem that under a universal law of asserting lies rather than the truth for self serving reasons would lead to a collapse of the institution of testimony.

According to the humanity formulation of the categorical imperative, we assess the moral worthiness of an action by considering whether it willing an action of that type involves respecting a person’s humanity rather than treating them as a mere means to an end.<sup>17</sup> Consider again the example of a shopkeeper not giving their customer fair change at the end during a business interaction. Intuitively, this kind of action does not

---

<sup>16</sup>Kant, 1785/2018 G 4:421. I use ‘G’ to indicate that I am citing from Kant’s *Groundwork for the Metaphysics of Morals*. I follow the conventions of Kant scholarship in using the academy numbers, e.g., n:n, for page citations.

<sup>17</sup>Kant (1785/2018) G :4:429

pass the humanity formulation of the categorical imperative. This kind of action does not pass the humanity formulation of the categorical imperative because it involves the agent treating customers merely as a means to make more money rather than respecting how the customer, as a human, has ends of their own that they would have used that fair change to pursue. Hence, according to the humanity formulation of the categorical imperative, the shopkeeper's action is not morally worthy. It is not morally worthy because it involves the shopkeeper's exploiting their customers in ways that do not respect or acknowledge how the customer has ends of their own that they would have used that fair change to pursue.

Similarly, consider the action of a speaker's asserting a lie rather than the truth for self-serving reasons. On Kant's view, such a speaker is using their audience as a mere means to their own ends rather than respecting their audience's humanity. Such a speaker is not respecting their audience's humanity because their audience, at some level, demands the truth for the pursuit of some aspect of their own life's projects.

Finally, consider the kingdom of ends formulation of the categorical imperative. According to this formulation of the categorical imperative, we assess the moral worthiness of an action by considering whether it is the kind of action that can be cogently passed as legislation in a merely possible kingdom of ends. Such a kingdom is a state composed of individuals who regard each other as ends rather than mere means and who legislate laws accordingly. Consider once more our example of the shopkeeper who does not give their customer fair change during a business exchange. Intuitively, this sort of action would not pass as legislation in a merely possible kingdom of ends. It would not pass because such a law would amount to a license for citizens of the kingdom to make exceptions of themselves by treating others as mere means. Such laws are antithetical to a possible kingdom of ends.

Now consider once more a speaker who, for self serving reasons, asserts lies rather



than truths. This speaker's action is not the sort of action that would pass the kingdom of ends formulation of the categorical imperative. This is not the kind of action that would pass as legislation in a possible kingdom of ends. Such a kind of action would not pass as legislation in such a kingdom because it would amount to a license for citizens of the kingdom to make exceptions of themselves by treating other citizens as mere means. Such laws are antithetical to a possible kingdom of ends.

All of Kant's treatments of lying out of self interest, with respect to the categorical imperative, point to the same conclusion. For Kant, lying out of self interest fails the categorical imperative. Such lies fail the categorical imperative because they essentially involve the liar making exceptions of themselves or imposing double standards upon others. Such double standards allegedly either lead to the breakdown of an associated institution when willed as a universal law, result in treating humans as mere means as opposed to respecting their humanity, or be antithetical to a kingdom of persons who respect each others humanity. And since, for Kant, the categorical imperative, is an ideal standard on rational willing, it would follow on Kant's view that a person who lies is not ideally rational.

I'm not saying that Kant's moral philosophy is the correct moral philosophy. However, I am saying that if my sketches of Kant are adequately accurate, then Kant's account of the moral status of lying is a precursor to a non-conciliatory response to the rational deception objection. Of course, it might be anachronistic to say that Kant was giving a non-conciliatory response. I am doubtful of whether ambitious presumptivism and the rational deception objection as I conceive them were ever on Kant's mind. Nevertheless, I think it's fair enough to say that Kant's account of lying avails, if not outright commits, Kant of a non-conciliatory response to the rational deception objection.

As far as I'm aware, Kant's account of the morality of lying has had little impact direct impact on contemporary, mainstream theorizing specifically about ambitious pre-

sumptivism.<sup>18</sup> Part of that is due to the fact that Kant’s discussions significantly pre-date those discussions. Part of it is due to the fact that Kant, unlike Thomas Reid, does not seem to be advancing anything like a predecessor to an ambitious presumptivist proposal. And part of that is also due to the harshness of Kant’s position. Kant’s account seems to deny the morality and rationality of all lies, even for seemingly good ends, such as altruistic and paternalistic lies.<sup>19</sup> For example, it is immoral, on Kant’s view to lie to would be murderers about the locations of their intended victims. Some think that Kant must be able to make exceptions in such cases, but there’s little agreement on how.<sup>20</sup> In light of these considerations, it’s unsurprising that Kant’s ethics goes unnoticed as a potential response to the rational deception objection.

## 1.5 Burge: The Contemporary Standard Bearer

Although Kant’s moral arguments for the non-conciliatory response haven’t played a major role in contemporary mainstream theorizing about ambitious presumptivism, at least one neo-Kantian take on the non-conciliatory response has. Tyler Burge offers an influential discussion of ambitious presumptivism in his paper “Content Preservation” (1993.) At first glance, a reader might not recognize it as a neo-Kantian argument because Burge’s arguments in that paper don’t rely on substantive theses about the relationship between morality and rationality. Nor do Burge’s arguments place any im-

---

<sup>18</sup>This isn’t to say that it hasn’t had any impact on broader philosophizing about testimony in general. In particular, it occupies a central place in the ethics and political philosophy of testimony. For a recent discussion of Kant’s place in those discussions, see Shiffrin (2014).

<sup>19</sup>Korsgaard (1996), p. 355

<sup>20</sup>Korsgaard (1996) offers an influential discussion, according to which the exception involves a kind of self-defense from exploitation towards cruel ends. However, as Shiffrin points out, there are plenty of cases like that of the inquiring would be murderer where it seems like it should be morally permissible to lie, but that at the same time it doesn’t seem as though self-defense isn’t a relevant concern ((2014), p. 31). Shiffrin offers an illuminating new alternative in terms of “suspended contexts” which purports to make up the difference, but it’s not yet clear what hold that proposal has, or will have, on mainstream philosophical discussion.

portant explanatory weight on other hallmarks of a Kantian account of morality, such as the categorical imperative. Instead, Burge’s discussion rests largely on a novel blend of theses about the nature of mental representation and the nature of naturally occurring function bearers.

Nevertheless, Burge, like Kant, offers a non-conciliatory response. Like Kant, Burge offers a novel argument according to which the falsity of RDO1 follows from a substantive defect in an agent’s capacity to “reason”. Where Kant locates the defect in a failure to satisfy crucial formulations of the categorical imperative, Burge locates the defect in a failure to satisfy a general unity requirement on reasoning. This unity requirement, together with Burge’s view that all reason, practical or theoretical, constitutively aims at truth, strikes me as a significant development for substantive theories of rationality.<sup>21</sup> I call Burge’s distinct take on the non-conciliatory response the “functional unity argument”.

## 1.6 Roadmaps

In the next chapter, I explain Burge’s neo-Kantian, functional unity argument for the non-conciliatory response. The remainder of the essay is marshaled towards defending both the functional unity argument and my attribution of it to Burge. I defend the viability of interpreting Burge as committed to the functional argument from those who do not read Burge as committed to a non-conciliatory response. I defend the functional unity argument from objections that its premises are too implausible, that it involves an equivocation between senses of rationality, and that it involves an equivocation between sense of “transpersonal” function. In defending the functional unity argument from that last objection, I forge a connection between Burge’s functional unity argu-

---

<sup>21</sup>As opposed to formal theories of rationality, such as the Bayesian account of rationality. For a philosophically influential statement of the Bayesian account of rationality, see Ramsey (Ramsey, 1926/2011).

ment and “truth-goal” and “reasoned discourse” accounts of assertion. Finally, I defend Burge’s functional unity argument from an objection in moral and political philosophy that dogged Kant’s moral philosophy. As a result, I reserve space for Burge’s functional unity argument and the emerging account of rational agency to make a contribution to the intersection of ethics and epistemology.

## Chapter 2

# The Functional Unity Argument

Tyler Burge's 'Content Preservation' is a landmark paper in the epistemologies of memory and testimony. There, Burge identifies the rational deception objection as a challenge for proponents of ambitious presumptivism. I believe Burge answers the challenge thus:

Reason necessarily has a teleological aspect, which can be understood through reflection on rational practice. Understanding the notion of reason in sufficient depth requires understanding its primary functions. One of reason's primary functions is that of presenting truth, independently of special personal interests. Lying is sometimes rational in the sense that it is in the liar's best interests. But lying occasions a disunity among functions of reason. It conflicts with one's reason's transpersonal function of presenting the truth, independently of special personal interests.

(Burge, 1993, 475).

I read this passage as committing Burge to a non-conciliatory response to the rational deception objection. The argument goes like this. First, reason is constitutively functional. Second, reason's constitutive function is composed of several sub-functions called 'aspects', including an aspect of presenting truths regardless of the reasoners wants. Third, another aspect of reason's constitutive function is that of promoting the reasoner's in-

terests. Fourth, reason's aspects are beholden to a unity requirement. Fifth, anytime a reasoner, as a speaker, asserts a lie rather than a truth, the speaker is disunified. Sixth, whenever a reasoner is disunified, they are not rational. Conclusion: anytime a reasoner, as a speaker, asserts a lie rather than a truth, they are not rational. In other words, it's not true that it's sometimes rational for a speaker to assert lies rather than truths. Let's explain each in turn.

## 2.1 Parsing the Functional Unity Argument

First, Burge is claiming that reason is constitutively functional. This means that reason, as an ability, is type individuated by its function, i.e., what it's "supposed" to do to promote an aim. Consider an analogy with hearts. Some people say that human hearts are constitutively functional, meaning that such hearts are type individuated by their function. In this case, human hearts are "supposed" to pump their body's blood to promote some level of blood circulation.

The heart analogy is an analogy with a *natural*, or *naturally occurring* function. Here's an analogy with an *artifactual* or *man made function*. Consider an analogy with personal computers. You might say that personal computers are also functional in the sense that they are type individuated by their design plans. In this case, personal computers are "supposed" to execute programs to promote their user's ends.

Plenty of ink has been spilled on the nature and fixation of functions, particularly natural functions.<sup>1</sup> I will stay neutral on those issues. I will instead help myself to a sketch of the *format* of function, be they natural or not, constitutive or not.<sup>2</sup> Consider the following function attributions:

---

<sup>1</sup>For some illuminating accounts of the nature and fixation of function that are in the vicinity of what Burge has in mind, see his (2009b), as well as Wright (1973), Millikan (1984), and Mossio *et. al* (2009).

<sup>2</sup>And by "help myself" I mean to deliberately beg the question against anyone who is "skeptical" about there being any genuine sense of functioning.

**FBA1** Hearts function to promote their host body's blood circulation

**FBA2** An individual's visual perceptual system functions to visually refer to particulars in the immediate environment

**FBA3** Knives function to cut

The trio FA1-FA3 are examples of function attributions. Each function attribution references an "actor" and an "end". The "actor" referenced is the function bearer. In the first example, the "actor" is the heart. In the second, it's the individual's visual system. In the third, it's knives.

The "end" or "goal" referenced in a function attribution specifies what conditions or events that the actor's behavior or functioning promotes. In example FA1, the end or goal is the promotion of a level of blood circulation. In FA2, the end is visual reference to environmental particulars. In FA3, the end is the activity of cutting. It is in the satisfaction of these goals that a function bearer's activity counts as successful.

When I say, for example, that hearts function to promote blood circulation by pumping blood, I am not only making general claims about how hearts behave. I am also describing hearts as beholden to certain performance standards. We must distinguish different dimensions for evaluating functional behavior. There's *successful* versus *failed* functioning. Successful functioning is when the function bearer's behavior promotes its aim. Hearts, for example, have the function of pumping blood in the service of maintaining some level of blood circulation. A heart functions successfully when its activity of pumping blood promotes certain levels of blood circulation. Otherwise, the heart functions unsuccessfully.

In addition to distinguishing successful and unsuccessful functioning, we should also distinguish normal functioning and malfunctioning. For this evaluation, we bracket away the behavior's relationship to its goal and we look at the behavior itself. If an occasion

of purposive behavior co-occurs with an “internal” breakdown, then it’s an occasion of malfunctioning. If not, then it’s an occasion of normal functioning. For example, if a heart’s pumping coincides with such internal breakdowns as tears in its chambers or its chambers pumping out of rhythm, that heart is malfunctioning. A heart whose activity isn’t associated with any breakdown is a heart that’s functioning normally.

Second, Burge is specifying an aspect of reason’s constitutive function. Sometimes, functions are composed of many “smaller” functions. The latter are “aspects” of the former. Burge is saying that an aspect of reason’s constitutive function is the aspect of promoting true presentations, regardless of the reasoner’s wants. This is often called the “theoretical” or “impersonal” aspect of reason.

In specifying the theoretical aspect of reason’s constitutive function, Burge specifies a functional goal. I think we can reasonably infer that Burge thinks that the reason includes belief forming processes such as inference and perceptual belief formation, and intention forming processes. Further, I think we can also say, given the context, the reason also includes the formation of assertions. After all, Burge formulates his version of ambitious presumptivism in terms of the reception of “presentations as true”. For Burge, testimony, i.e., assertions, fall under this category.

Third, Burge is articulating another aspect of reason’s constitutive function. This much is clear when he says that “Lying is sometimes rational *in the sense that it is in the liar’s best interests*”.<sup>3</sup> Burge thinks that another aspect of reason’s constitutive function is that of promoting (the reasoner’s) desire satisfaction. When we talk of rationality in terms of doing promoting your desires satisfaction, we’re referencing this specific aspect of of reason’s constitutive function. This is often called the “instrumental” aspect of reason.

Again, in specifying the instrumental aspect of reason, Burge specifies the aspect’s

---

<sup>3</sup>My emphasis.



goal but not its means. And again, we can deploy some ordinary intuitions to fill the gap. I think we can reasonably infer that Burge thinks the functional means includes capacities for actions mediated by plans or intentions. This includes capacities for forming plans/intentions and for executing the behaviors called for by such plans.

Fourth, Burge is invoking a “unity” requirement on functioning. I believe Burge’s unity requirement is a standard for *normal* functioning. It is most illuminating in contexts where the function in question consists of two or more aspects. A function is unified when and only when all its aspects are being promoted together. We might state the unity requirement thus: if something functions *normally*, then its function is unified. This means that whenever something’s function is disunified, then it is malfunctioning. A vacuous case of unified functioning is when a capacity’s function cannot be meaningfully decomposed into smaller aspects.

I think reflection on intuitions reveals a rationale for linking disunified function with malfunctioning. Consider the case of a dysfunctioning music band. Music bands function to play music.<sup>4</sup> Members must play their parts *together*. A drummer might be an individual virtuoso and nevertheless overpower their bandmates’ contributions. If so, then even if that drummer individually plays drums well, their band overall plays poorly. The band would be disunified because some players would be pursuing their own excellence at the expense of the other players. Accordingly, this band would be dysfunctional because they are disunified.

By the same token, I think that these reflections also support linking unified functioning with functioning normally. Consider the case of a functional music band. As before, we’re assuming that music bands function to play music. And we’re again helping ourselves to the assumption that band members must play their parts *together*. Intuitively,

---

<sup>4</sup>Of course, I’m assuming that it makes sense to think of some social groups as functional, and I’m assuming that it makes sense that bands are among such social groups.

a music band that is unified in their playing, i.e., a band that is playing together, will be overall excellent. For example, if the drummer is individually the best musician in the band, then the drummer playing together with the rest of the band involves setting a beat of such a tempo and complexity that the other band members can follow and find that their parts receive the correct accentuation at the correct time. At the same time, the other players following along with the drummer's playing adds harmonic and melodic content on top of the drummers beats. Even if the individual performances are nothing special, the whole performance is excellent when the band performs unified.

What goes for the normal functioning of a band of musicians also goes for the normal functioning of an individual rational agent. We might think of an individual rational agent as someone who must promote the instrumental aspect of reason and the impersonal aspect of reason. Intuitively, a person who promotes the instrumental aspect of reason at the expense of the impersonal aspect could get lucky and succeed in their actions. For example, such a person might be more successful in escaping captivity if they believe, against their evidence, that it's probable that they will escape their captors.

Nevertheless, it seems that a person whose aspects of reason are unified generally does better than a person whose aspects of reason are not unified. For example, a person whose practical reasoning is grounded in their total body of successful impersonal reasonings, i.e., in their total stock of evidence, is a person whose practical reasoning seems to enjoy a greater robustness against failure. Different accounts of ideal rationality parse this out differently. For example, proponents of knowledge first accounts of ideal rationality seem to maintain knowledge is the standard for impersonal rationality.<sup>5</sup> On this view, the value of being impersonally rational derives from the value of knowledge. The value of knowledge, in turn, is that it provides some level of "safety" from falsity and failure. Accordingly, just as an impersonally rational person's whose beliefs are impersonally

---

<sup>5</sup>C.f., Williamson (2000)

rational cannot easily be false, an impersonally rational person's actions cannot easily fail.

Another account of ideal rationality that suggests the superiority of unified reasoning to disunified reasoning is "probabilism". Probabilism is the view that a person's impersonally rational beliefs are representable as probabilities. Traditionally, probabilism has been supported by so called "dutch book" arguments.<sup>6</sup> Setting aside the sophisticated mathematics used to formulate such arguments, the alleged point of those arguments is that a person who is not impersonally rational, at least in the sense that their beliefs are representable as probabilities, is a person who is unable to guard against self-defeating actions. Again, the broader idea being suggested here is that a person whose practical and impersonal aspects of reasoning are unified is a person who enjoys some kind of protection from practical failure.

Fifth, Burge is claiming that lying to others is itself an occasion of reason's being disunified. Specifically, he thinks that an occasion of a speaker's asserting a lie is an occasion where speaker elevates the instrumental aspect of promoting the reasoner's self-interest at the expense of the impersonal aspect of presenting the truth regardless of self-interest. Given the unity requirement, a person who lies pits reason's aspects against each other on that occasion. The person is disunified on that occasion.

We can lend plausibility to Burge's claim that lying to others is another occasion of reason's being disunified by considering assertion's role at the interpersonal level of social reasoning. The interpersonal aspects of assertion's role in social reasoning are most clearly exhibited by reflection on assertion's role in the "game of giving and asking for reasons". To bring this out, I'm going to help myself to some aspects of a dialectical account of assertion and I'm going to use those aspects to recount the essentials of Rescorla (2007)'s under-appreciated explanation of how such an account of assertion provides a deeper

---

<sup>6</sup>Ramsey (1926/2011)

ideal explanation of why it seems that speakers only assert what they sincerely know.

On a dialectical account, assertion is individuated by its role in the game of giving and asking for reasons<sup>7</sup>. This “game” is often called “reasoned discourse”<sup>8</sup>. Reasoned persuasion is a game where two or more players work to rationally persuade the other of their idiosyncratic view. Reasoned inquiry is a game where two or more players work together to marshal reasons to settle a question of truth. Assertion essentially involves making a commitment to that proposition’s truth. That commitment is constituted by its association with two norms:<sup>9</sup>

**Defense** When challenged to defend an asserted proposition, one must provide a non-circular, cogent argument for the proposition or else retract it.

**Retraction** When faced with a counterargument against an asserted proposition, one must rebut the counterargument or else retract the proposition.

In asserting a proposition, you undertake responsibility for arguing for its truth when challenged or faced with counterargument. When you can’t give those arguments, you must retract—i.e. cancel—your commitment and thereby relinquish your responsibility. “Challenges” may be understood as raising any doubts or questions as to a proposition’s truth. This is the “asking for reasons” part of rational dialectic. Counterargument is what you’d think: someone has “given reasons” against your assertion. On this view, you’re licensed to assert whatever you like. However, once you make an assertion, you’re accountable for defending it or taking it back.

Rescorla’s formulation enhances the dialectical account with constitutive *goals* of rational dialectic:<sup>10</sup>

<sup>7</sup>See Brandom (1994), MacFarlane (2003) and Walton and Krabbe (1995) .

<sup>8</sup>Rescorla (2007) p. 254

<sup>9</sup>Rescorla (2007) points out that Brandom (1994) and MacFarlane (2003) endorse a restricted version of the Defense Norm.

<sup>10</sup>Rescorla, 2007 p. 255

**Rational Rapprochement** Rational dialectic has the function of rationally isolating relevant, mutually acceptable premises.

This is a constitutive goal of rational dialectic because speakers who don't strive towards it cease to engage with each other rationally. Schoolyard name-calling conflicts are not rational engagements insofar as the participants don't strive towards **Rational Rapprochement**. Two corollary constitutive goals guide interlocutors towards **Rational Rapprochement**:<sup>11</sup>

**Avoid Losing** Avoid counter-arguments against what one asserts that are decisive against oneself.

**Try Winning** Provide arguments in favor of what one asserts that are decisive against one's opponent.

Let's say that an argument is decisive against someone just in case they can't marshal reasons against it. **Avoid Losing** captures the intuitive idea that you're not promoting **Rational Rapprochement** if you continue to be stymied by counter-arguments. **Try Winning** captures the idea that you are promoting **Rational Rapprochement** if your opponent must accept your assertions on the strength of your arguments. All together, these three goals exert pressure against overly casual asserting and retracting. If you're always retracting your assertions at first challenge, you're contravening either **Avoid Losing** or **Try Winning**. Flouting either of those goals means that you're not striving towards **Rational Rapprochement**. If you're not striving towards **Rational Rapprochement**, then you're either playing poorly, or not playing at all.

Here's my tentative reconstruction of Rescorla's argument:

### No Asymmetries Argument

<sup>11</sup>The labels are my own. Rescorla leaves these unlabelled.

**NA** Absent relevant evidence, you are practically entitled to presume no epistemic or cognitive advantage over your interlocutor.

**NAA1** Absent relevant evidence, if you are practically entitled to presume no epistemic or cognitive advantage over your interlocutor, then either stonewalling, dissembling (i.e., asserting lies), or truthfulness (i.e., asserting truths) is the rational default dialectical strategy.

**NAA2** Absent relevant evidence, stonewalling is not the rational default dialectical strategy because it contravenes **Rational Rapprochement**

**NAA3** Absent relevant evidence, dissembling is not the rational default dialectical strategy because dissembling is more costly than truthfulness given **NA** and **Avoid Losing**.

**RD** Absent relevant evidence, truthfulness is the practically rational default dialectical strategy.

The argument's main steps involve modus ponens and disjunctive syllogism. This means that the argument is deductively valid. **NA** is compelling as a reflection on the default set up of an idealized game of reasoned discourse. In everyday, ordinary life realizations of games of giving and asking for reasons, the players are rarely, if ever, on even footing with one another. Some players are cleverer than others while other players know more about the relevant topics than others. In some contingencies, there will invariably be players who are able to assert lies probably get away with it. Now contrast ordinary real life contingencies with the ideal conditions isolated by the goals and rules of a game of giving and asking for reasons. In such conditions, players know nothing about each other. They do not know whom among them is better at argumentation. They also do not know who among them is the superior expert on the subject matter at hand.

Accordingly, assuming they're good faith players, i.e., players who participate with the goal of winning, the ideal practically rational default is for players to assume that they possess no relevant epistemic or cognitive advantage over other players.

**NAA1** is compelling because it captures the basic strategies available in a rational dialectic. Stonewalling corresponds to asking a lot of questions without offering any assertions of your own. Dissembling is asserting falsehoods in non-circular defense of an asserted proposition. Truthfulness is asserting truths in a non-circular defense of an asserted proposition. In practice, mixed strategies exist. However, those mixed strategies seem to have some structure. You might assert deceitfully in order to motivate the central questions in your stonewalling strategy. You might assert truths in order to conceal your dissemblings. One of the basic strategies will be the endgame to a mixed strategy.

**NAA2** is also compelling. Insofar as the goal is to come to a rational agreement, it's unhelpful to offer nothing while continuously challenging your interlocutor's contributions. This results in you and your interlocutors walk away without agreeing on anything.

**NAA3** leverages comparative considerations against dissembling. Given **Rational Rapprochement**, you have to make some unretracted assertions. Given **Defense** and **Retraction**, your assertions foist upon you commitments to provide non-circular arguments. Given **Avoid Losing**, you need to avoid decisive counter-arguments. Given **NA**, it's rational to take yourself as the baseline for how your arguments would be received. Dissembling involves asserting, and so defending, assertions you believe false. Dissembling is more costly in part because it involve making an elaborate network of lies.<sup>12</sup> This means that dissembling involves the construction of complex fabrications. In contrast, its considerably less resource intensive to assert propositions you believe to be true.

The costliness of dissembling also derives from its inherent riskiness. You have no

---

<sup>12</sup>Rescorla (2007) p. 266

control over whether and when the truth will get revealed.<sup>13</sup> And the more complex your fabrications—i.e. the more lies you’ve had to spin in defense of your earlier lies—the greater the risk posed to you by the revelation of truth. After all, the revelation of truth amounts to a decisive counter-argument to your lie. The more your assertions turn upon a lie, the greater the risk that your assertions face decisive counter-argument when that lie is exposed. In contrast, the revelation of truth aides truthful asserters in marshalling decisive counter-arguments against their opponents. Whereas dissembling increases the risk of violating **Avoid Losing**, truthfulness increases the ease of fulfilling **Try Winning**. Absent relevant evidence, truthfulness is the optimal strategy for satisfying the goals of rational dialectic. As such, it is the rational default dialectical strategy.

Rescorla’s **No Asymmetries Argument** suggests a plausible way to illustrate the link between asserting the truth and being unified and between asserting lies and being disunified. The link is this: A speaker who is unified in the sense that Burge has in mind is a speaker who follows the ideally rational default dialectical strategy. As such, they would enjoy the protection normally afforded to those who speak the truth in an idealized game of giving and asking for reasons. In contrast, a speaker who lies—a speaker who is disunified in Burge’s sense—is one who is in principle incurring a systematically riskier dialectical strategy. It is one that is not as plausibly going to achieve rational rapprochement because it is in principle and systematically more vulnerable to decisive objections.

You might have doubts about the relevance of Rescorla’s argument to Burge’s functional unity argument. You might have such doubts because you might think that Burge’s and Rescorla’s arguments don’t fit well together. On the one hand, Rescorla’s goals are more modest. Rescorla doesn’t want to assume anything like a constitutive role for truth telling in our practices of reasoned discourse. On the other hand, Burge seems happy

---

<sup>13</sup>Rescorla (2007) p. 266



to assume a constitutive role for truth telling in our form of agency. Given these differences, it is fair to wonder exactly how relevant or congenial their considerations are to each other.

For right now, I maintain that the connection to draw is a very modest one. When explaining the import of Burge's unity requirement, I talked about how individual reasoners whose reason was unified enjoyed a systematic advantage or benefit over those whose reason was disunified against the impersonal aspect of reason. To help illustrate the plausibility of Burge's extension of the impersonal aspect of reason, and hence the unity requirement, to assertion and testimony, I have appealed to Rescorla's explanation of why we ought to speak the truth to explain how a speaker who is unified, in Burge's sense, enjoys a systematic advantage over a speaker who is disunified.

I am not saying, at this point, that Rescorla's arguments and Burge's arguments fit together neatly. However, I am saying that Rescorla's argument illustrates, in ways that might be helpful later, that just as individuals who are impersonally rational in their beliefs enjoy a systematic advantage to those who are not, so too does a person who is impersonally rational in their assertions enjoys a systematic advantage to those who are not impersonally rational in their assertions. In Rescorla's hands, the systematic advantage enjoyed by an impersonally rational speaker is some minimum level of either protection from challenge or counterexample, or else some minimum guarantee of being able to competently address such challenges or counterexamples. Given his account of reasoned discourse, this advantage amounts to a baseline level of protection from failing to achieve rational *rapprochement*.

For Burge, there are at least two ways to assimilate the lessons of Rescorla's argument. He could insist, as I am insisting for him, that Rescorla's argument illuminates a broader point about rationality and reasoned discourse. That broader point, again, is that impersonally rational people enjoy a systematic, baseline advantage over those

who are not impersonally rational. Alternatively, Burge could go further by leveraging something analogous to Rescorla's argument as part of a defense of his assumption that truth telling plays a constitutive role in our form of agency. For now, I will sit content with the first route and return to that second route in chapter five of this essay.

Up until now, I have been defending the fifth premise of Burge's functional unity argument. That premise says that a person's lying to others is itself an occasion of a person's reason being disunified. I now turn elaborating upon the sixth premise of that argument.

As I understand it, in the sixth premise, Burge is claiming that a reasoner who is disunified is not rational. Here's reason to think that Burge links an agent's being disunified with that agent's not being rational:

Reason has a function in providing guidance to truth, in presenting and promoting truth without regard to individual interest. This is why epistemic reasons are not relativized to a person or to a desire. It is why someone whose reasoning is distorted by self-deception is in a significant way irrational-even when the self-deception serves the individual's interests.

Burge (1993) p. 475

Here, Burge is talking about how reason's function of presenting truth, regardless of individual interests, bears on the explanation of the irrationality of self-deception. A self-deceiver may believe against their evidence, because it's more convenient for them. Nevertheless, in doing so, their reasons are promoting their self interests at the expense of presenting truth regardless of self-interests. They are irrational because they are disunified.

To illustrate what underlies this line of thought, consider the individual accounts of reasoning that I have recounted while discussing the fourth and fifth considerations

of Burge's functional unity argument. These were Williamson's knowledge-first account of rationality, the traditional Dutch Book Arguments for Probabilism, and Rescorla's No-Asymmetries Argument for why sincerity is the ideally rational default dialectical strategy. When we zoom out from the obvious differences between these arguments and focus on the common core, a certain pattern emerges. In each of those discussions, the core is this: a person whose reason unifies the instrumental and the impersonal aspects of reason enjoys a systematic advantage or benefit versus someone whose reason is not so unified. Specifically, even if a person might, on some occasion, luck out while favoring the practical aspect of reason over the impersonal aspect of reason, a person who unifies both aspects enjoys a systematic advantage over those who do not.

So, the pattern is that people whose reason is unified enjoy a systematic advantage over people whose reason is disunified in favoring the practical aspect of reason over the impersonal aspect. However, it's not just the existence of a systematic advantage that makes unified reasoners rational and disunified reasoners irrational. It's also the nature of that systematic advantage that matters. The advantage posited in each of the previous examples is a type of protection from risk. According to a knowledge first rationality, the protection consists in a person's plan's being based in beliefs that couldn't easily be false. According to the Dutch Book Arguments for probabilism, the protection consists in immunity to certain forms of guaranteed losses. According to Rescorla's No-Asymmetries Argument, the protection consists in minimizing one's exposure to decisive counterexamples and hence preserving an optimal ratio of successful defenses to issued retractions.

Generalizing on those discussions, we can say something in favor of Burge's sorting of unified reasoners with the rational reasoners and disunified reasoners with irrational reasoners. The reason why it's plausible to link unity to rationality and disunity to irrationality is that those who are unified enjoy systematic protections from certain types

of risks that those who are disunified do not enjoy. It is this exposure to risk that constitutes the significant way that a person who is disunified is irrational.

Here's an additional, independent intuition pump in favor of saying a disunified reasoner is not rational. Burge's unity requirement on functioning is analogous to consistency requirements on instrumental rationality. Recall that on instrumental rationality, or as Burge might call it, the instrumental aspect of reason, the focus is on promoting a reasoner's desire satisfaction by forming actions mediated by plans. There are three ways a plan can be inconsistent. I think the unity requirement may be independently supported as a generalization of one of these consistency requirements.

A plan can be "means-ends" inconsistent. Such a plan is one where the settled means either undermines or does not contribute toward the settled end. Self-defeating strategies are examples of such plans. If I want to foster friendship with someone, and I know that verbally abusing them does the opposite of that, then I would be irrational in verbally abusing them for the sake of fostering friendship.

A plan can also be "means-means" inconsistent. A plan with inconsistent means is a plan where one of the component means towards the end goal undermines or conflicts with another component means towards the same goal. Knowingly endorsing such a plan is instrumentally irrational. This suggests that rationality requires "means-means" consistency.

A plan can also be "ends-ends" inconsistent. There can be an inconsistency between a plan's specified ends. In such a case, two or more of a plan's component ends cannot be pursued together. For example, a person's plans might include as ends being monogamously related to one person and being monogamously related to another. Unfortunately, monogamy doesn't work that way; you can only be monogamously related to one person and no one else. Pursuing one person comes at the expense of the other. A person who knowingly endorses a plan with such incompatible goals is being instrumentally irrational.

The unity requirement finds a clear analogue in the rational requirement for “ends-ends” consistency. “Ends-ends” consistency violations involve scenarios where a person knowing adopts plans where they have two ends but must sacrifice one end for the other. Insofar as a person’s plan mediated actions function to promote their ends, we might say that a violation in “ends-ends” consistency occasions a disunity in their action’s function because one of their ends is pursued at the expense of another.

Burge’s unity requirement seems most analogous to the requirement for “ends-ends consistency”. The analogy rests in how failures to satisfy the requirement involve behaviors that favor one end at the expense of the other. The difference, if there is one, is a difference of generalization. The “ends-ends” consistency requirement seems to be a unity requirement restricted to the case of plans; i.e., a person’s representationally mediated commitment to act for certain ends.

It doesn’t strike me as too implausible to think of plans as relatively transient, revisable functions that are taken on by rational agents. Korsgaard (2009) defends such a proposal. Korsgaard thinks that being a rational agent involves, among other things, unifying your seemingly disparate ends under a plan that promotes all of them. If this is plausible, then the unity requirement is a generalization of the “ends-ends” requirement by abstracting from features peculiar to rational agency and focusing instead on what’s common to all complex functioning.

I can now extrapolate how Burge is committed to a non-conciliatory response to the rational deception objection. When a person asserts a lie rather than a truth, they’re disunified. They’re disunified because in lying, they’re favoring reason’s aspect of promoting self-interest at the expense of the aspect of presenting truths regardless of self-interest. And when they’re disunified, they’re irrational. They’re irrational because a reasoner’s being disunified suffices for them to be irrational. So, whenever a person asserts a lie rather than the truth, they’re irrational. Therefore, it’s not true that it is sometimes

rational for a person to assert lies rather than truths. So concludes the functional unity argument.

The functional unity argument is a species of non-conciliatory response to the rational deception objection. This is so because it contradicts **RDO1**, which states that it is sometimes rational for a speaker to assert lies rather than truths. Since Burge is committed to the functional unity argument, he is committed to a non-conciliatory response to the rational deception objection.

## 2.2 Is The Functional Unity Argument Nonempirical?

For Burge, a claim is *a priori* when it can be nonempirically justified. A justification, i.e., a mode by which a claim is justified, is nonempirical when it doesn't rely on empirical considerations. You might question whether the functional unity argument I am attributing to Burge is nonempirical. The question matters because Burge believes ambitious presumptivism is *a priori*. This means that my interpretation of Burge would be implausible if the functional unity argument is empirical. I address that worry here, in this subsection.

I grant Burge's demarcation between empirical and nonempirical considerations.<sup>14</sup> For Burge, empirical considerations essentially reference the specific contents of specific sensations, sensory perceptions, perception based beliefs, and memorial recollections. Nonempirical considerations do not. Burge thinks that nonempirical justifications proceed by reflecting upon a person's own understandings of concepts, principles, the general structure of their experiences. Burge takes inspiration for his account of the empiri-

---

<sup>14</sup>After all, I want to give his proposal as fair a chance as I can.

cal/nonempirical distinction, and the emergent *a priori/a posteriori* distinction, in his reading of Kant.<sup>15</sup>

Let's build on Burge's demarcation by also distinguishing a justification's being apparently nonempirical from it's being ultimately nonempirical. A justification is *prima facie* nonempirical when it apparently doesn't deploy or assume empirical premises. A justification is ultimately nonempirical when it actually doesn't deploy or assume empirical premises. Many people advance apparently nonempirical arguments. When their apparently nonempirical arguments fail to be ultimately nonempirical, it's because they unwittingly relied on empirical assumptions. Those assumptions are often so banal, deeply ingrained, or obscure that we don't notice that they're empirical.<sup>16</sup>

I don't think my interpretation fails if the functional unity argument isn't *ultima facie* nonempirical. For my purposes of charitable interpretation, I need only establish how the functional unity argument is, at least by Burge's lights, *prima facie* nonempirical. And I think that the functional unity argument is indeed *prima facie* nonempirical. Here's how.

First, there's Burge's claim that reason is constitutively functional. I think here we can charitably say, on Burge's behalf, that this claim is somehow *a priori*. When people think about whether reason is constitutively functional, they can reflect upon their understanding of what reasoning is and upon what role that reasoning plays in their lives. They appreciate how reasoning seems to have essentially succeeded when a reasoner arrives at certain sorts of attitudes in a certain manner and how reasoning seems to have essentially failed when a reasoner arrives at other sorts of attitudes in other sorts of manners. The apparent fact that we essentially evaluate reasoning in terms of success or failure to meet a goal is enough, it seems, to give us a *prima facie* nonempirical

---

<sup>15</sup>Burge, (2000).

<sup>16</sup>(Burge, 2010), pp. 534-537

rationale for thinking that reason is constitutively functional.

Second, there's Burge's claim that reason's constitutive function is broken down into smaller aspects, one of which is the aspect of presenting truths regardless of the reasoner's interests. Again, I think we can charitably say, on Burge's behalf, that this claim is also *a priori*. A person may reflect upon a more specific role that reasoning plays in their lives, such as the role of theoretical reasoning. Upon reflection, theoretical reasoning is essentially evaluated in terms of the goal of arriving at true beliefs about the world. It is successful when it does, unsuccessful when it doesn't. Further, theoretical reasoning is beholden to standards of internal integrity: we can scrutinize an instance of theoretical reasoning for whether it manifests structural properties such as inductive coherence and deductive validity. Those standards would correspond to what it is for theoretical reasoning to function normally, i.e., without internal breakdown.

Third, there's Burge's claim that another aspect of reason's constitutive function is that of promoting the reasoner's best interests. I think we can charitably say, on Burge's behalf, that this claim is *a priori*. This goes similarly to the sketch offered in the previous paragraph. A person may reflect upon a more specific role that reasoning plays in their lives, i.e., the role of practical reasoning. Upon reflection, practical reasoning is essentially evaluated in terms of the goal of settling upon plans for promoting the reasoner's desires and life projects. Practical reasoning is successful when a person does settle upon such plans, and unsuccessful when they don't. Further, practical reasoning is beholden to standards of internal integrity: we can scrutinize an instance of practical reasoning for whether it manifests structural properties such as means-ends coherence, means-means coherence, and ends-ends coherence. Such standards would correspond to what it is for practical reason to function normally, i.e., without internal breakdown.<sup>17</sup>

---

<sup>17</sup>For a proposal closer to Burge's own novel proposals, consider his *a priori* discussion of primitive agency in his (2009b). In that discussion, he reflects on the role of teleological explanations in our practice of explaining behavior and argues that the most basic form of agency, primitive agency, involves



Fourth, there's Burge's imposition of a unity requirement on the function of reason. Here, we can charitably say on Burge's behalf that this claim is *a priori*. The reflections I offered earlier, regarding the distinctions between the different dimensions of assessment for a function, and regarding how the notion of unity could help explain malfunction in an arbitrary case, were *prima facie a priori*. They involved coming to an understanding of the sorts of distinction we can make when we evaluate a functional performance. And they involve understanding how an arbitrary function can be composed of several other functions.

Further *a priori* support for a unity requirement is suggested by Burge's terse discussion of self-deception. As a brief refresher, he writes:

Reason has a function in providing guidance to truth, in presenting and promoting truth without regard to individual interest. This is why epistemic reasons are not relativized to a person or to a desire. It is why someone whose reasoning is distorted by self-deception is in a significant way irrational—even when the self-deception serves the individual's interests.

Burge (1993) p. 475

Self deception, at a bear minimum, involves attempting to believe and to act on the supposition of the truth of proposition that you, on some level, know or have good evidence to believe to be false. The idea is that when we reflect upon how we understand an arbitrary, idealized case of self-deception, it seems to us that the case consists in disunity. That discussion didn't depend on the particulars of a single actual case of self-deception. Instead, we brought to bear our understanding of what self-deception is and of what the function of reason is, and we unpacked how those understandings applied to

---

the production of behaviors in the service of goals. At the level of primitive agency, the agent need not, and often cannot, represent anything at all. And nevertheless, their actions are evaluated in terms of the goals they promote.

that arbitrary case.

Fifth, there's Burge's claim that lying to others is itself, an occasion of reason's being disunified. This claim follows from *prima facie a priori* reflection on Burge's previous three claims. Since those claims are *prima facie a priori*, Burge's fifth claim is also *prima facie a priori*.

You might worry that my defense of Burge's fifth claim, specifically the discussion of Rescorla's No-Asymmetries argument, introduces considerations that are empirical in nature. This worry is warranted insofar as that argument is offered as part of the philosophical literature on assertion. Participants in that literature sometimes view the debate as an empirical one. They might think this because arguments about which theory of assertion is superior to another rely upon claims about which theory handles "the data" better than the other. For example, in his classic discussion on theories of assertion, Williamson frames the debate as one where we frame a minimal "hypothesis" about assertion and then "test" it against certain cases.<sup>18</sup> Nevertheless, for reasons I touch upon in chapter five, section two of this essay, I am not worried. Even if the structure of the debate is as Williamson says it is, the initial "tests" philosophers conduct aren't genuinely empirical tests. There aren't observational or controlled studies of populations. Instead, these tests encapsulate thought experiments concerning our non-empirical understanding of the practice of reasoned discourse. While there is certainly room in this methodology for observational and controlled studies, those empirical considerations seem to serve as tie breakers between explanatorily competitive proposals that already enjoy a substantive level of *a priori* warrant.

Sixth, there's Burge's claim that a reasoner who is disunified is not rational. His defense of this claim is based on *prima facie a priori* reflection on an understanding of what it is to be a self-deceiver. That understanding is supplemented by *prima facie a*

---

<sup>18</sup>Williamson (2000) pp. 238-239

*priori* reflections on the constitutive functional aspects of reason. Since both supports are *prima facie a priori*, so too is Burge's claim that a reasoner who is disunified is not rational.

In light of the foregoing, I claim that all six premises of the functional unity argument are *prima facie a priori*. This means that, absent the unwitting presumption of an empirical premise, Burge's claims commit him to the further *prima facie a priori* claim that it's not true that it's sometimes rational for a speaker to assert lies rather than truths.

# Chapter 3

## Objections and Replies, Round One

In the previous section, I explained why I think it's plausible to interpret Burge as committed to the functional unity argument for rejecting RDO2. In this section, I consider the objection that my interpretation is implausible.<sup>1</sup> By replying to these objections, I strengthen the case for reading Burge as I do. I arrange these objections so that the strongest is at the end.

### 3.1 Objection One: Rejection-Out-of-Hand

Some people might object to my interpretation by simply rejecting it out of hand. It often the first response I encounter when I tell people that I read Burge this way. This response is often motivated by some combination of the following intuitions:

**ROH1** Rejecting RDO1 is incoherent because RDO1 is clearly obvious

**ROH2** The functional unity argument is so trivially flawed that someone as smart as Burge wouldn't argue for it

**ROH3** Burge didn't intend to advance the functional unity argument

---

<sup>1</sup>Plenty of people have said that they think that my interpretation of Burge isn't merely implausible, but obviously wrong. Almost all of them have said so without argument. In this section, I assume that they have one, and that it's one of the objections discussed. I suppose I might be too charitable in doing so.

People who feel some combination of these ways tend to feel that my reading of Burge is simply too implausible on its face to merit serious consideration.

## 3.2 Reply to Objection One

I think this is the weakest objection to reading Burge as I do. Rejecting something out of hand is sometimes warranted. But at other times, it's simply question begging. Here's why I think rejecting my interpretation out of hand is simply question begging.

First, RDO1 is a philosophical thesis. This is so no matter how intuitive or obvious it seems. Our all things considered grounds for believing RDO1 will ultimately turn on getting a deeper, philosophical understanding of what it is to be rational. Among other things, this chapter advances that dialectic. I think this speaks against ROH1's import.

Second, the functional unity argument is a sophisticated argument. It deploys substantive philosophical reflections on the nature of functioning. It may turn out that these reflections fail. But as philosophers, we must work to explain how. Among other things, this chapter does just that. I think this speaks against ROH2's import.

Third, it doesn't follow from the fact that a person doesn't intend something that they're not committed to it. For example, Frege didn't intend to give a logically inconsistent account of arithmetic. Still, Frege's work committed him to such an account. Similarly, Burge may not have intended to advance the functional unity argument. Still, his own words might commit him to it, at least on that occasion. I think this speaks against ROH3's import.

### 3.3 Objection Two: “Burge Believes RDO1”

Some people might object to my reading of Burge because they believe that Burge maintains that it’s sometimes rational for a person to assert lies rather than truths. Indeed, he seems to say as much. I quote:

Lying is sometimes rational in the sense that it is in the liar’s best interests.

Burge (1993) p. 475

People might take this quote to mean that Burge grants that lying is sometimes rational. That by itself isn’t enough to generate a disagreement between their reading of Burge and mine. To get there, my opponents might deploy either of two additional intuitions.

The first intuition is that Burge is saying that it is sometimes on-balance rational for a person to assert lies rather than truths. A person’s action is on-balance rational when their total stock of reasons supports their performing that action. This stands in contrast with *pro tanto* rationality. A person’s action is *pro tanto* rational when their total stock of reasons might fall short of supporting their performing that action. So, the intuition is that Burge is saying that it is sometimes rational for a person to assert lies rather than truths, where the person’s reasons are enough for their action to be rational.

The second intuition is that Burge is not presenting unity as a rational requirement. Recall that when a person’s action is disunified, it manifests their promoting some aspects of reason at the expense of the rest. When we say that unity is not a rational requirement, we’re saying that the person’s action being disunified doesn’t entail that their action is irrational.

These two intuitions together entail a reading on which Burge grants, if not outright believes, RDO1. Here is how. The first intuition means that, in some situations, a person can have enough reason to rationalize lying. The second intuition means that a person’s rationality cannot be undermined or overridden by disunity. It means that we

can transition from a person's being instrumentally, *ultima facie* on-balance rational in lying, to their being *ultima facie* on-balance rational *simpliciter* in lying. And if that's what Burge has in mind, then my reading of Burge as committed to the functional unity argument is incorrect.

At least one prominent commentator appears to read Burge the way described above. Consider how Graham discusses the case of self-deception:

Suppose a wife has plenty of evidence that her husband is having an affair, evidence that would justify the suspicions of a conscientious detective. But suppose that if the wife believed the evidence, she would lose all faith and the marriage would be ruined. Suppose furthermore that her life would be ruined if the marriage collapsed. In such a case, she might successfully self-deceive herself into believing that her husband is not having an affair but is simply overworked. She might then be all-things-considered better off, even though she is "in a significant way irrational" for she has crossed the theoretical dimension of rationality.

Graham (2018) p. 107.

Graham reads this as a case of on-balance, rational self-deception. Accordingly, the wife's self-deception seems to be sufficiently rationalized by how thoroughly her life would be ruined if she believed in accord with her evidence. And it seems as though Graham doesn't think that the ensuing disunity undermines or overrides the on-balance rationality of her self-deceiving.

Here's further evidence for reading Graham this way. Commenting further on the self-deception case, he writes:

Burge allows that you can be all-things-considered rational in any of these cases. Even so, you have "crossed rationality in one significant dimension—

the impersonal theoretical dimension.” Even if all things considered it can be rational to believe against the evidence, believing against the evidence is still a failure in “generic” rationality; it can be all-things-considered rational to fall short of epistemic rationality. For reason as a whole to work well, these functions must not be at cross-purposes... Full generic rationality involves a unity in the functioning of these two dimensions of rationality, and I agree. (Graham, 2018), p. 108.

In this passage, Graham grants that Burg is correct about “generic” rationality. Graham also denies that “generic” rationality is the decisive standard for a person to be on-balance rational. Instead, according to Graham, a person can be on-balance rational in either of the two ways. A person could be impersonally rational and not practically rational, e.g., this person is doing their best to believe truths and avoid falsity, but perhaps some of their plans or intentions don’t really add up in the “practically rational” way. Accordingly, this person has done the bare minimum that rationality demands of them and they enjoy the status of being on-balance rational.

Alternatively, a person could be practically rational and not impersonally rational. Such a person’s plans measure up in the “practically rational way”, but they’re not doing well in terms of believing truths and avoiding falsehoods. According to Graham, this person has satisfied the bare minimum for being fully rational.

Now consider the person who is both practically and impersonally rational. This person’s plans and intentions measure up in the “practically rational” way. This person is also doing well enough believing truths and disbelieving falsehoods. According to the view of rationality Graham is offering, this person has gone beyond the bare minimum demanded for on-balance rationality. This person is “generically” rational. But, since generic rationality goes beyond the minimum demanded for on-balance rationality, it fol-



lows that generic rationality is merely superlative. This is how Graham seems to conceive of the relationship between the different aspects of rationality, on-balance rationality, and generic rationality. And as far as I can tell, Graham seems to think that Burge would also understand rationality in this way.

Graham's way of reading Burge might be further elucidated by appeal to something like a distinction between ideal and non-ideal accounts of rationality. When we offer ideal accounts of rationality, we're trying to give something like a definition or explanation of the concept of rationality that would generally apply here, there, and everywhere. As such, we abstract from the particularities of individual circumstances and we maybe even help ourselves to assumptions that are unrealistic. Such assumptions are called "idealizations".

For example, Newton's physics gives an ideal account of gravity. It is an ideal account of gravity because it assumes the existence of frictionless planes. There are no such planes in the real world. Similarly, Bayesian accounts of practical rationality, such as the account by Savage (1954), is an ideal account of practical rationality. It is ideal because it helps itself to the assumption that a person is able to process arbitrarily many representations or arbitrary complexity. In reality, this assumption seems false. No person has the bandwidth nor the processing power to actually process arbitrarily many representations of arbitrary complexity. Their mental bandwidth is limited and they cannot process certain representations quickly enough.

You might think that as it is with Newton's physics and Bayesian accounts of practical rationality, so it is with Burge's account of generic rationality. Just as Newton's physics assumes planes that do not really exist and just as Bayesian accounts of practical rationality attribute to people powers they don't actually have, so too does Burge's account of generic rationality help itself to certain idealizations. You might then supplement this line of thought by maintaining that when we talk about what rationality requires of

people in real world circumstances, they're being ideally rational is superlative after all. You might think this because for some reason you worry that a person's achieving the state of being unified is an unrealistic idealization. All that rationality requires of real beings is that they be rational either practically or impersonally.

### 3.4 Reply to Objection Two

I now describe the best reply to the objection that Burge believes that it is sometimes rational for a person to assert lies rather than truths. My reply is that I resist the two intuitions that I think support the reading that Burge believes RDO1. My reply is a limited one; I do not aspire for it to be decisive against someone who genuinely holds those intuitions. After all, in the face of recalcitrant information, a person might plead that something else has gone wrong. Still, I can at least aspire to explain how someone may be able to resist those intuitions. I'll discuss each in order.

The more innocent of the two intuitions is the first one. Recall that the first intuition is that, when Burge says lying is sometimes rational, he is saying that it is on-balance rational. On the one hand, it doesn't seem necessarily so. For example, it might be that Burge is simply saying that lying is *pro tanto* rational. That is to say, lying enjoys sometimes enjoys *some* rational support because lying, at the very least sometimes satisfies the liar's interests. And we might say that the occasional *pro tanto* rationality still falls short of being on-balance rational because it is outweighed by how it engenders a failure of unity. So, I need not concede that Burge meant to say that lying is sometimes on-balance rational.

On the other hand, I don't think that I need to reject this specific intuition outright. I could concede that Burge is saying that lying is occasionally on-balance rational, and add this occasional on-balance rationality is merely *prima facie*. On this reading, the

role of Burge's subsequent claim that lying engenders a disunity among reasons functions is to flag that there's a systematic defeater to what would otherwise have been adequate support for lying. This is, I think, broadly compatible with my attributing the functional unity argument to Burge. On this reading, the functional unity argument articulates the structure of our systematic defeat for the occasional on-balance rationality of asserting lies rather than truths. So I need not necessarily reject the first intuition.

The dialectically stronger of the two intuitions is the second one. The second intuition said that, when Burge presents the unity requirement, Burge presents a supererogatory level of performance rather than a necessary and sufficient requirement for baseline rationality. I don't have a silver bullet objection to this intuition, and I won't aspire to give one. Instead, I offer a battery of considerations that together, speak against the plausibility of this intuition.

I begin by reproducing Burge's discussion of the self-deception case once more:

Reason has a function in providing guidance to truth, in presenting and promoting truth without regard to individual interest...It is why someone whose reasoning is *distorted* by self-deception is in a significant way irrational-even when the self-deception *serves the individual's interests*.

Burge (1993) 475, emphases added

Let's call this the "self-deception passage". As I read this passage, the impersonal, theoretical aspect of promoting truth is what explains how a self-deceiver's reasons are *distorted*. If unity were a superlative or supererogatory requirement, it would be inappropriate to describe failures of unity as distortions. To describe something as a distortion is to describe it as relevantly defective. Something is relevantly defective when it suffers a relevant flaw or inadequacy.

Now consider what it is for something to be flawed or inadequate. Flaws and inad-

equacies are relative to standards. Something is inadequate when it falls short of the relevant standard. For example, when we evaluate flaws in a copy, as a copy, we're looking to see if the copy has properties that diverge from the original. In this case, the original sets the standards relative to which the we evaluate the adequacy of the copy. This follows from the goal of copying. The baseline standard for a successful copy is maximal, relevant qualitative sameness between an original and its facsimile. Similarly, if a distortion in reason entails that there's a relevantly significant way in which someone is irrational, then it seems that the person has fallen short of the minimal standards for baseline rationality.

I think it's plausible to read Burge as describing disunity as a flaw or a failure, not as mere difference. I think this because irrationality isn't a merely descriptive property, like redness or circularity. It marks a flaw, or failure. And it seems plausible to me, given the passage, that the distortion involved with self-deception is a failure in being rational. Now, I might be wrong. Maybe Burge doesn't think that self-deception involves a relevant failure in rationality. But it doesn't read that way. If Burge didn't think that self-deception involved a distortion, he wouldn't have described it as such. Further, if Burge didn't think that distortion correlated with a failure for baseline rationality, he wouldn't have described it as irrational. Burge describes self-deception as a distortion, and he correlates that distortion with irrationality. This, to me, seems to be Burge describing self-deception as a failure to meet the standards of rationality, not as Burge describing an odd way of meeting those standards.

Proponents of the intuition that Burge's unity requirement will reply that I'm begging the question against the intuition. They will claim that Burge is describing a transgression in theoretical rationality, not rationality *simpliciter*. And rationality *simpliciter* would be this mongrel class consisting of two independent forms of rationality. To satisfy the minimal baseline of rationality *simpliciter* would be to satisfy at least one of these

forms. And to be supererogatorily rational *simpliciter* would be to rational in both ways.

I don't think that the immediately preceding rejoinder is plausible. Burge, in the self-deception passage, doesn't explicitly describe the self-deceiver as rational, *simpliciter* or otherwise. He merely describes them as having served their own individual interests while being "in a significant way irrational". I think it's more natural to read this as Burge saying that such a person is irrational despite having satisfied their own interests. I think it's natural to read him this way because, if he meant to say that the liar was rational, *simpliciter* or otherwise, he would have said so. He doesn't say so.

Further, I offer a different route to understanding "ways" in the self-deception passage. A proponent of the intuition that Burge's unity requirement is supererogatory reads "way" as referencing a "sense", as though there were fully independent senses of the term 'rationality'. These distinct senses would impose independent conditions upon the world. This would support the view that rationality *simpliciter* is a mongrel class composed of independent standards. If you follow that reading, then my reading of Burge's passage isn't very plausible.

And yet, that's not the only way to read "way". Sometimes, we use talk of ways to describe, not different senses of a polysemous term, but different manners in which the conditions determined by a single sense go satisfied. Suppose I ask an all powerful genie to make me the healthiest person in the world. There are at least two manners in which the genie could satisfy my demands. The genie could leave everyone else's healthiness as it is and increase my healthiness beyond what has currently been achieved by other people. Alternatively, the genie could leave my health as it is and reduce the healthiness of everyone else in the world so that my current level of healthiness would be the highest level currently achieved. Either constitutes a "significant way" of being the healthiest person in the world. But they don't necessarily constitute different meanings.

Similarly, we might say deceiving oneself is a significant way of failing to satisfy the

condition of being rational. And we might say that failing to protect one's interests is another significant way of failing to satisfy that same condition of being rational. And finally, we might say that that condition of being rational is determined by the univocal meaning of being rational. Read this way, the point of the unity requirement is to bridge what seem like two disparate aspects of rationality into a single category: rationality *simpliciter*. I think that this way of understanding Burge's self-deception passage is no less plausible than the reading on which Burge is switching around between two independent senses of "rational". At the very least, the intuition that the unity requirement is not a minimal standard is optional. I don't have to endorse that intuition or bring it to bear when reading Burge's self-deception passage.

Finally, I am not convinced of the efficacy of invoking the distinction between ideal and non-ideal rationality to make plausible the idea that being unified, that being generically rational in Burge's sense, is a supererogatory level of performance. One reason I'm not convinced is that it makes what threat rational deception poses to ambitious presumptivism meaningless. Essentially, all of these people who have spilled ink into making both premises of the rational deception objection plausible shouldn't have bothered. For ambitious presumptivism is a thesis about the epistemic dimension of what's ideally rational for a person to do. Accordingly, pointing out that it's non-ideally on-balance rational for a person to assert lies rather than truths is neither here nor there.

Another reason for why I am not convinced of the efficacy of invoking the ideal/non-ideal rationality distinction is that the idealizations that made Newtonian physics and Bayesian decision theory ideal theories are different from Burge's purported unity requirement for generic rationality. Contrary to Newtonian physics, there are no frictionless planes. Contrary to Bayesian decision theory, there are propositions too big for a person to have rational beliefs about. These are clearly idealizations. In contrast, the claim that a person must unify their reasons is not like these idealizations. A person's reasons

being unified is not a counterfactual scenario. People really do unify their reasons all the time. For example, the scientist who publishes that a certain research paradigm has failed, even if that's the paradigm for which they got funding, has unified their reasons. The thief who turns themselves in and admits the truth of their crimes in exchange for a lighter criminal sentence has also unified their reasons. The person who accepts the evidence of their limitations, even if it there was the hope that it would be easier for them to perform had they ignored that evidence, has unified their reasons. A person being unified in their reasons is difficult. But it can and does occur.

## Chapter 4

# Objections and Replies, Round Two

In the first section of this paper, I explained why I think it's plausible that Burge is committed to the functional unity argument. I sketched what I thought the premises of the argument were and why I thought they comprised a seemingly nonempirical argument.

In the second section, I considered some objections according to which I'm mistaken in attributing the functional unity argument to Burge in the first place. There, I gave my reasons for resisting the intuitions against attributing the functional unity argument. I did not, however, consider objections to the substance of the functional unity argument itself.

In this section, I consider and reply to three objections to the substance of the functional unity argument. These objections are motivated by uneasiness with different facets of Burge's claims about the functionality and unity of reason. Obviously, people's uneasiness isn't a thing that can be directly rationally engaged with. However, by engaging with the different ways this uneasiness manifests in philosophical commentary, I think that the uneasiness's impact can be reduced. I have arranged these objections so that the strongest of the three is discussed last.



## 4.1 Objection Three: The “Non-Starter” Objection

One *prima facie* compelling objection is that the functional unity argument isn't *a priori* because the unity requirement is a non-starter. The rough idea is that the *a priori* of some of its premises is too remote to warrant serious discussion or elaboration. It is roughly comparable to how someone like G.E. Moore might reject the *a priori* possibility of skeptical hypotheses in favor of their claim to know they have hands on the power of perception. Granted, not everyone is enamored with this kind of thinking. And among those who are, there's controversy about how it might work. Nevertheless, I propose to set aside the mechanics of this sort of response and instead to focus on why someone might favor it.

A recent statement of the non-starter objection is found in Joseph Shieber's thorough (2015) review of the epistemology of testimony. For the most part, Shieber offers an insightful reconstruction of the overall line of argument Burge offers for Burge's formulation of ambitious presumptivism. However, when it comes to Burge's response to the rational deception objection, he stops. Shieber rightly acknowledges the role of a unity requirement in Burge's response. Nevertheless he dismisses it as too *prima facie* implausible to warrant serious discussion.

Shieber's reasons for dismissing Burge's response are succinctly stated. He writes:

If Burge is to appeal to the unity of reason, it would seem that he must demonstrate that all of our rational norms—including those involved in practical rationality—are directed at truth. While it seems much more promising to argue that norms of theoretical rationality are united in aiming at truth, it would certainly seem difficult to motivate the idea that all of our norms are truth directed.

Shieber (2015,) p. 118

At the heart of the dismissal is a reading of Burge’s discussion on which, because of the unity requirement, both theoretical and practical rationality must aim at promoting truth. But, for Schieber at least, it seems overwhelmingly plausible that practical rationality is aimed at desire satisfaction rather than truth. Since Schieber grants theoretical rationality’s truth directedness, he rejects the unity requirement.

## 4.2 Reply to Objection Three

Here is my take on the structure of Schieber’s objection:

**SHO1** If it’s *a priori* that an aspect of reason is that of presenting truth and it’s *a priori* that reason is subject to a unity requirement, then it’s *a priori* that all of reason’s aspects are “truth directed”

**SHO2** It’s not *a priori* that all of reason’s aspects are “truth directed” because practical reason isn’t so.

**SHO3** It’s *a priori* that an aspect of reason is that of presenting truth

**SHO4** So, it’s not *a priori* that reason is subject to a unity requirement

Schieber spends no time elaborating SHO1. Neither will I, since I can instead try attacking either of SHO2 or SHO3. Since Schieber grants Burge SHO3, so I won’t attack that either. Instead, I’ll attack SHO2.

The justification for SHO2 is that practical reason is not “truth directed”. Someone might think this because they think a reasoner’s self-interests, i.e., their desires, plans, or goals, aren’t the sorts of things that are “truth-directed”. Let’s contrast, for example, two different attitudes a person may take towards the proposition that their cat is healthy. A person might *believe* that their cat is healthy. We could scrutinize a person’s belief that their cat is healthy on the grounds that their cat is, in fact, not healthy. That’s because

beliefs are “truth-directed” in the sense that they’re evaluated on whether they present the world as it is.

Alternatively, a person might *desire* that their cat be healthy. We do not scrutinize a person’s desire the same way we scrutinize their beliefs. If a person’s cat isn’t healthy, we don’t necessarily scrutinize them for desire that their cat be healthy. Indeed, we might even scrutinize them for not having a desire that their cat be healthy. This strongly suggests that person’s desire isn’t necessarily defective for not presenting the world as it is. Accordingly, this suggests that a person’s desires aren’t “truth-directed”.

Similarly, a person may *intend* or *plan* that their cat be healthy.<sup>1</sup> Again, we often do not scrutinize a person’s plans the way that we scrutinize their beliefs. The current unhealthiness of a person’s cat is compatible with their planning that their cat be healthy. This is so because such a plan might take some time to complete. Overall, it seems as though desires and plans aren’t the sorts of things that are “truth-directed”.

The above reflections indicate some fairly intuitive ways in which states of practical reason, e.g., desires and intention, are not “truth-directed”. Nevertheless, I contend that those are not the only substantive ways for practical reason to be “truth-directed”. There are at least two other substantive ways in which practical reason for practical reason to be “truth-directed”. The first is that truth-directedness and “fulfillment” are species of a broader, truth oriented category. The second is that successful practical reasoning pre-supposes successfully truth-directed premises.

Let’s start with that first way in which a reasoner’s desires and plans are “truth-

---

<sup>1</sup>I tend to talk about intentions and plans interchangeably. I am not endorsing so-called “planning” theories of intention, according to which intentions not reducible to special cases of beliefs, desires, or combos of beliefs and desires. Nor am I endorsing the idea that in addition to possible worlds, which are used to explain a belief’s satisfaction conditions, there are also “plans”, which are used to help explain an intention’s satisfaction conditions. Instead, I think I’m latching on to a folk concept that is expressed equally well by the words ‘intend’ and ‘plan’. Any theory of intention, including the “planning” theory, has as its job the articulation of that folk concept. For those who wonder why I bring up this disclaimer at all, see the development of the planning theory in Bratman’s (1987). For a recent compendium of critical discussions, see Vargas and Yaffe (2014).

directed”. Like a belief, a person’s desires and plans are essentially associated with *satisfaction conditions*. A mental state’s satisfaction conditions, be it a belief, desire, or a plan, are determined by its content. A belief’s satisfaction conditions are the conditions in which the belief’s content are accurate. A desire’s satisfaction conditions are the conditions in which the desire is fulfilled.<sup>2</sup>

The difference between accuracy and fulfillment corresponds to two different senses of the phrase made true. Accurate beliefs are always made true by how the world is. At minimum, beliefs are accurate when they are in accord with the world. Sated desires are, at least sometimes, made true by the desirer.

The distinction between accuracy and fulfillment is even sharper when comparing beliefs and plans. A person’s plan that their cat be healthy is made true in those circumstances where the person saw to their plan’s *execution*. We may call a plan’s satisfaction conditions its *execution* conditions. When a person’s desires are made true as a consequence of their executing their plans, those desires are made true by the person themselves.

Arguably, the transformation of desires and goals into plans is within practical rationality’s purview. I think this is so because practical rationality essentially involves the capacity of practical reasoning, and practical reasoning involves the transformation of desires and goals into plans. This is analogous to how theoretical rationality essentially involves the capacity of theoretical reasoning, and how theoretical reasoning involves the transition from settled beliefs to additional beliefs.

The above discussion suggests two important different ways in which theoretical and practical rationality are truth-directed. Theoretical rationality aims at the truth in the sense of an individual’s “reflecting” the truth. By this, I mean that an individual’s beliefs are defective if they fail to accord with the actual facts. In contrast, instrumental rationality aims at the truth, if at all, in the sense of spurring an agent to seeing to it

---

<sup>2</sup>Searle (1983), Elizondo (2013) p. 4, Rescorla (2015)

that their plans are made true. Put plainly, instrumental rationality aims at the truth in the sense of aiming agents into causing things to be true.

I contend that it is *prima facie a priori* that practical rationality is truth-directed in the sense indicated in the last paragraph. It is *prima facie a priori* because we arrived at this claim through seemingly nonempirical means. Those means were reflection upon our seemingly nonempirical understanding of the structure and contents of our own mental states and reflection on our nonempirical understanding of how our mental states are evaluated.

Now let's consider the second substantive way in which a person's practical reason is truth-directed. The second way in which practical reason is truth-directed is that successful occasions of practical reason are constrained by successfully truth-directed premises. We might say that practical reason is, in principle, *indirectly* truth-directed.

I don't need to appeal to Burge's account of rationality to make sense of how successful practical reasoning requires successfully truth-directed premises. I think intuitive reflection on good cases of practical reasoning should make the case well enough. Suppose that you have a desire to have pork sisig for dinner. What is it for that desire to be fulfilled? Intuitively, it is for it to be true that you have pork sisig for dinner.

The mere having of a desire does not make for a complete occasion of practical reason. To be part of an occasion of practical reason, a person's desire must stand in the right relations to certain other thoughts of theirs so as to constitute an intention or plan with fairly determinate execution conditions. That is to say, a person's desire only really counts as part of practical reason when it partly constitutes that person's plans, i.e., their attitudes for *making true*.

What is it for your desire to have pork sisig for dinner to constitute an intention to make true that you have pork sisig for dinner? It's like this. First, you need to appreciate your different ways of knowing how to get pork sisig for dinner. These include cooking

pork sisig for yourself. These also include ordering it from a Filipino restaurant. Second, you need to know which of those different ways you know how to get pork sisig for dinner fit into your situation. If you're too far away from a Filipino restaurant, you're left with the option of cooking it yourself. And if you don't know what constitutes cooking pork sisig, your desire to have pork sisig for dinner cannot enter into a genuine intention to make true the proposition that you have pork sisig for dinner.

The above example strongly supports the idea that it's not just that practical reasons are truth-directed in the sense that practical reasons are attitudes of making-true. They are also truth-directed in that genuine occasions of practical reason are partly constituted by successfully truth-directed premises. Suppose that your practical premises do not include genuinely accurate ways to acquire pork sisig and that they do not accurately reflect which of those ways are feasible given your situation. In that scenario, your plan to have pork sisig for dinner is, in some way or other, a failure of practical reason. It could be a failure of practical reason because you might be mistaken about accurate ways to acquire pork sisig and which of those ways is feasible given your situation. In that case, practical reason failed because your theoretical reasoning failed. It could instead be that you suffer a failure of practical reason because, despite your evidence decisively telling you otherwise, you still plan to have pork sisig for dinner. In that case, practical reason failed because you fail to appreciate your theoretical reasoning.

The above reflections seem to me to be *prima facie a priori*. They strike me as such because it's not like we need to now formulate a study where we observe some people practically reason in order to test the hypothesis. The role of the example is to exercise our own concepts of practical reason and to see what makes sense. Thus, these reflections are *prima facie a priori* considerations.

Both sets of the preceding reflections together weave a picture where practical reason is multiply truth-directed. Practical reasoning is truth-directed in the sense of directing

people to make propositions true. It is also truth directed in the sense that it cannot be successful without drawing upon a person's stock of knowledge and warranted beliefs.

Finally, I assert that my *prima facie a priori* claims about the truth directedness of practical rationality are largely, if not entirely, congenial to Burge's overall philosophical programme. It seems to me that Burge distinguishes rational agency from other forms of agency, in part because rational agency involves more than centrally guided functional behavior towards goals, and in part because rational agency involves more than centrally guided functional behavior towards perceptually represented goals. Rational agency also requires centrally guided towards conceptually represented goals. Arguably, a person conceptually represents their goals, if at all, by forming beliefs about their goals and by forming plans about how to accomplish those goals.

### 4.3 Objection Four: “Equivocating” on the Senses of Rationality

Another objection is that Burge's claims about the aspects and unity of reason involves an “equivocation” between two senses of rationality. Elizabeth Fricker advances such an objection in her classic (2006). As we will see, it seems more accurate to call the alleged problem a “gap” rather than an “equivocation”. However, I follow Fricker's diction out of respect for the original source material.

Fricker interprets Burge as attempting to underwrite the following defense of ambitious presumptivism:<sup>3</sup>

If you in some manner receive and apprehend what seems to be a message that M, an attempt at assertoric communication of information, then:

---

<sup>3</sup>Fricker (2006, p. 75)

- i you can presume it really is such a message; hence
- ii you can presume the sender is rational; hence
- iii you can presume she is a “source of truth”.

Let’s grant the veracity of this interpretation. According to Fricker, the conflation occurs in Burge’s defense of the transition from (ii) to (iii). As she sees it, we can grant Burge a “thin” sense of rationality in which “an event’s being rational does indeed entail that its agent is rational”.<sup>4</sup> Fricker thinks we can grant Burge this thin sense of rationality when defending (ii).

This thin sense of rationality that Fricker has in mind is the Bayesian account of rationality. As a proposal about theoretical rationality, the Bayesian prescribes that a person believe propositions in proportion to the support the proposition enjoys from the person’s total stock of evidence. As a proposal about practical rationality, the Bayesian prescribes that a person take the course of action that given their total evidence, yields the highest average payoff.

The Bayesian account of rationality is often invoked by philosophers of mind, such as Donald Davidson, who think that mental states, specifically propositional attitudes such as beliefs, desires, and plans, are constitutively governed by rational standards.<sup>5</sup> For such philosophers, this is so because Bayesian decision theory, together with a suitably sophisticated formal semantics, gives us the tools to formulate meaningful empirical theories of action, mind, and linguistic meaning.<sup>6</sup> The explanatory power of these theories, in turn, provides further support for the claim that mental states are constitutively governed by

---

<sup>4</sup>Fricker (2006) p. 77

<sup>5</sup>See, for instance, the essays in Davidson’s (1984), (Davidson, 2004), and his (2005).

<sup>6</sup>Of course, this is not a straightforward, easy to apply recipe for a theory of mind. Famously, Davidson’s deployment of these ideas had the result that there were no principled explanatory relationships in psychology, and hence no real psychological science be had. Davidson seems to have thought this was okay. You yourself might not. For illuminating critical discussions connecting Davidson’s idiosyncratic positions on psychology to the broader the philosophy of science, see chapter 4 of Woodward, 2003 and chapter 7 of Burge, 2010.



rational standards.<sup>7</sup>

This “Davidsonian” application of the Bayesian account is relevant for two reasons. First, Fricker references it when she introduces the thin sense of rationality. Second, Davidson’s proposal has had a lasting impact on the philosophy of mind. Philosophy of mind is taken for granted when explaining people’s assertions and other linguistic actions. Philosophy of mind has largely moved on to evaluating the classical computational theory of mind. On that view, the mind is constitutively a computing machine.<sup>8</sup> Within this paradigm, folk psychology is taken as the starting point to be regimented in light of empirical study. Within this paradigm, theories of rational standards still play important roles in formulating theories of mind and behavior. Some philosophers argue that in idealized conditions the mind’s computational processes approximately conform to Bayesian standards for beliefs and desires.<sup>9</sup> This suggests that the mind might be a Bayesian computing machine. This also suggests that explanations of what people do with words might be regimented in broadly Bayesian terms.

As Fricker says, this Bayesian account of rationality is compatible with the soundness of the rational deception objection. I think Burge would probably agree with this assessment. Fricker’s alleged trouble for Burge seemingly lies in the transition from (ii) to (iii). That transition is directly threatened by the rational deception objection. After all, the rational deception objection denies that we may presume, even provisionally, that a person is a source of truth because it’s occasionally rational for them to assert lies rather than truths.

Fricker claims that Burge’s response to the rational deception objection is to invoke

---

<sup>7</sup>If this sounds circular, that’s because it probably is. Davidson unashamedly maintains that justification, of any sort, involves some degree of circularity and cohesion. Indeed, he thinks that this circularity is required for there to be any semantic content at all. For more, see his (1986), where he argues that this circularity provides the beginnings of a transcendental argument against external world skepticism.

<sup>8</sup>See, for instance, Fodor (1975) and Gallistel and King (2011).

<sup>9</sup>For a recent proposal along these lines, see Oaksford and Chater (2007).

a rich sense of rationality. This rich sense of rationality is essentially the picture of rationality presented in my reconstruction of Burge’s functional unity argument, and which was developed further in my reply to Shieber. On this view, reason is essentially teleological, the constitutive function of reason is comprised of several truth-directed aspects, and those aspects are governed by a unity requirement on normal function. We might call this rich sense of rationality the Veridicality Account of Rationality. Fricker claims that Burge’s invocation of a Veridicality Account constitutes an “equivocation” between the thin and rich senses of rationality. Put plainly, what Fricker is actually doing is pointing out a gap between the thin and rich senses of rationality. Accordingly, when Burge gives what I’ve called the functional unity argument, he is purportedly failing to appreciate how the difference between the thin and rich senses of rationality entails a logical gap between (ii) and (iii). That is to say, her objection is that it doesn’t follow from the presumption that a sender is rational that a sender is a “source of truth”.

Of course, you might wonder how Burge might have made such a simple blunder. On this question, Fricker speculates that Burge is engaged in an “explanatory” rather than a “suasive” justification of the acceptance principle. Roughly speaking, a “suasive” justification is a non-question-begging argument that persuades skeptics to abandon their skeptical ambitions. In contrast, an “explanatory” justification is meant to reassure non-skeptics of their position’s plausibility. Allegedly, Burge might not have noticed the mistake because he allegedly didn’t set out to persuade skeptics with non-question-begging arguments.<sup>10</sup>

---

<sup>10</sup>Fricker offers this objection, that Burge conflates “thin” and “rich” senses of rationality, as a long preamble to another, more famous objection. Fricker’s more famous objection is that Burge’s *a priori* defense of ambitious presumptivism serves a thesis that’s irrelevant to explaining ordinary behavior. I think Burge gives an insightful reply to this objection in the postscript to “Content Preservation”, where he emphasizes that ambitious presumptivism:

...indicates that the minimum source of warrant for receiving communication is more general than human social context. The source lies in something universal to intelligible, propositional presentations-as-true (centrally, assertions). That universality, together with the way dialectical reasoning works with possible defeaters of default entitlements, shows

## 4.4 Reply to Objection Four

I think the beginnings of a reasonable reply on Burge’s behalf is the reply that he actually gives. In a footnote in the postscript to “Content Preservation”, Burge insists that Fricker’s objection:

...relies on the claim that in discussing rational lying, I equivocate between a ‘thin’ sense of being rational that is constitutive of having propositional attitudes and a presumably thicker sense: ‘wholly impartial and disinterested speaker only of the truth’, 82. Fricker thinks that my account of the Acceptance Principle depends on conflating these “senses” of being rational. I deny that two such senses of being rational are at issue. I make no use of her supposed thicker sense. I deny that the sense of being rational that is implicit in having propositional attitudes (and being capable of making assertions) is thin. I think that it is the full sense of being generically rational, as distinguished from being critically rational.

Burge (2013) p. 266

There’s a lot going on in this passage, but there are some key takeaways. The central takeaway is that Burge himself denies having made any equivocation in the first place. Instead, he places the sense of rationality that he has in mind as a middle ground between the Bayesian Account and a “wholly impartial and disinterested truth seeker” account. What I think he means by this is that, on the Veridicality account of rationality, reason

---

the source to be in a sense basic. All further considerations are more specialized, even if they are more common. (2013, p. 268)

As I understand him, Burge is interested in giving a principled account of a minimal, context-invariant aspect of our justified testimony based beliefs. This minimal, context-invariant facet of justified testimony based beliefs often gets obscured in every day life because of how it interfaces with particularities of context. But even when that minimal, context-invariant facet is hidden, it still plays a role in structuring how the different particularities of context interact and combine with each other to modify, positively or negatively, the justification supplied by the initial minimal context invariant facet.

isn't simply interested in accuracy for accuracy's sake. Rather, accuracy matters because it is in the service of facilitating the reasoner's attitudes for making truth.<sup>11</sup>

Another takeaway is that, even though Burge might agree with Davidson that propositional attitudes, are individuated by rational standards, he isn't committed to saying that those standards are accounted for by the Bayesian account. Instead, he can commit to the Veridicality account as a theory of generic, baseline rationality.<sup>12</sup>

While I think Burge's reply is a good start, I want to improve upon it. I want to do so because it's not a satisfactory endpoint. One reason why it's not is that his reply seems to rest on the idea that Fricker's wasn't talking about his account of rationality at all. And that's fine, so far as it goes. But Fricker doesn't actually say what she thinks Burge's "thick" sense of rationality" is supposed to be until the end of the discussion where she recapitulates the argument with a slogan. I think that Fricker can reasonably reply that she simply misspoke or that she was giving a quick caricature, and that the spirit of what she said applies equally well to his preferred "thick" sense of rationality. She can make that additional claim because in his reply, Burge doesn't marshal any textual evidence to dispel the charge of equivocation. He simply elaborates on his preferred sense of reason and rationality. When Burge elaborates on his preferred sense of reason and rationality,

---

<sup>11</sup>This focus fits in with Burge's discussion of the origins of representational agency (Burge (2009b)). The idea is this. The most basic form of agency precedes all sorts of representation. This is primitive agency. It is simply goal oriented behavior. The next level of agency is representational agency. This is behavior oriented around a *perceptually represented* goal. The idea is that well-functioning of such agency is constrained by the accuracy of the mediating perceptions. If those perceptions are not accurate enough, such agency results in failure.

<sup>12</sup>Burge references critical rationality; as I understand it, critical rationality is a more demanding set of capacities and standards that involve the ability to represent mental states as mental states, rational standards as rational standards, and the ability to scrutinize whether mental states (represented as mental states), conform to rational standards (represented as rational standards). Let's call "critical rationality" the set of rational standards that are essentially associated with critical reasoning. It is unclear to me whether critical rationality, as Burge understands it, comprises a set of supererogatory standards for rationality rather than a set of merely erogatory standards for creatures with a different set of capacities. Luckily, I don't have to address that issue now because the current issue is about the nature of generic, baseline rationality.

he emphasizes points that are introduced as part of the functional unity argument.<sup>13</sup> This ignores the facet of Fricker's objection on which it is the functional unity argument itself where the fallacious reasoning occurs. A decisive refutation of Fricker's objection needs to not only elaborate upon the account of reason Burge deploys, but also to give evidence that that account of reason was operative at the very beginning of Burge's discussion of ambitious presumptivism.

So, let's do it better. Here's what I would add to Burge's reply that no such conflation between "thin" and "rich" senses of rationality occur. Burge can and should insist he is merely elaborating on one and the same sense of rationality for the justifications of what Fricker has labelled (ii) and (iii). When providing Fricker's reconstruction of Burge's justification of ambitious presumptivism, in which (ii) and (iii) occur, Fricker cites a passage at Burge (1993, p. 472). Prior to the cited passage in support of (ii), there's already an indications that Burge has something in mind that's "richer" than a Bayesian account of rationality.

I believe the strongest indication that Burge, in "Content Preservation", has always been relying on something "richer" than a Bayesian account of rationality is when he writes:

if something is a rational source, it is a prima facie source of truth. For a condition on reasons, rationality, and reason is that they be guides to truth... An epistemic reason for believing something would not count as such if it did not provide some reasonable support for accepting it as *true*... If one has a reason or entitlement to accept something because it is, prima facie, rationally supported, one has a reason or entitlement to accept it as true. A source is a guide to truth *in* being rational.

---

<sup>13</sup>Specifically, Burge repeats his claims that lying essentially involves a "transgression" by the practical aspect of reason against the theoretical impersonal aspect of reason, and that this transgression means that the person falls short of being fully generically rational.

Burge(1993) p. 470-471

Again, the passage the Fricker cites in support of (ii) occurs at p. 472. The passage quoted above, where Burge begins to explain what he thinks is rationality, occurs before that, on p. 470. This very passage ends at page 471. In this passage, the idea is that a rational source, whatever that is, is the sort of thing that exercises reason (a functional capacity) to produce reasons (products of said functional capacity) that settle questions about what to do or think, while conforming to rational norms (standards for exercising reason successfully and standards for exercising reason normally). Burge is claiming that truth promotion or “guidance towards truth” is a condition on something’s being a reason *qua* capacity, reason *qua* product, or a rational norm. It doesn’t matter whether Burge meant this to be a sufficient or a necessary condition because either way this condition outstrips a Bayesian account of rationality.

Truth promotion is not a sufficient condition for a belief’s being Bayesian-rational because a person’s having true beliefs doesn’t entail that their beliefs are probabilistically coherent. Probabilistic coherence is a matter of the strength of a belief, not its veridicality. A person might have a true beliefs, but the strength of such beliefs may be disproportionate to what’s probable on their evidence or be probabilistically incoherent. For example, my evidence might provide positive support for thinking that I may attain some desired job, without entailing that it is a full certainty. If my reaction to my evidence is to believe, with full certainty, that I will attain the desired job, then the strength of my belief is disproportionate to what’s probable on my evidence. I would not be rational in the thin, Bayesian sense

More importantly, truth promotion is not a necessary condition for a belief’s being Bayesian rational because none of a person’s beliefs need be true for their beliefs to be probabilistically coherent. A person’s beliefs may be a morass of false conspiracy theories.

But they might, nevertheless, hold those beliefs in such a way that which the whole mass could be representable as probabilities. If so, such a person would have beliefs that would satisfy the demands made by Bayesian norms. But none of those person's beliefs would count as "guides towards truth".

I submit that it's implausible that Burge commits any such conflation between a "thin" sense of rationality and a "rich" sense of rationality. Accordingly, contrary to Fricker's arguments, there's no "gap" between (ii) and (iii). I say so because truth-promotion is neither necessary nor sufficient for Bayesian rationality, because truth-promotion is central to the "richer" Veridicality account of rationality, and because Burge opens his initial defense of the ambitious presumptivism by appealing to truth-promotion. Those considerations tell me that, even if Burge's Veridicality account of Rationality is compatible with the Bayesian account of rationality, it is the Veridicality account, not the Bayesian account, that has been operative from the start. I.e., Burge's arguments for ambitious presumptivism, up to and including the functional unity argument, invoke a "thick", truth directed sense of rationality. If that's so, then Fricker is simply mistaken about the occurrence of such a mistake

## 4.5 Round Two Debrief

In this section, I considered two objections to Burge's claims about the nature of rationality. Both of these objections, I argue, can be handled by elaborating further on material that's already present in Burge's discussion. There's nothing wrong with saying that rationality is about truth promotion if we distinguish between two different modes of promoting truth: accuracy versus execution. Accuracy promotion is about "being true". Execution promotion is about "making true". Instrumental rationality is to making true as theoretical rationality is to being true.

Further, Burge doesn't occasion any equivocation between a thin, Bayesian account of rationality and his richer Veridicality account of rationality. Burge's discussion makes clear that he was using a veridicality account of rationality from the start. The ideas Burge invokes when introducing his Veridicality account are neither necessary nor sufficient for satisfying the demands of the Bayesian account. This isn't to say that they're not compatible. Rather, it's to say that there's no plausible way to see Burge as deploying or assuming anything like the Bayesian account.



# Chapter 5

## Objections and Replies, Round Three

In this section, I consider what I think is the most *prima facie* compelling objection to Burge’s functional unity argument that I can still dispel on Burge’s behalf. This objection is powerful enough that it cannot be dispelled by further developing Burge’s theoretical commitments. Instead, I dispel the objection by drawing connections between the functional unity argument and what I call “truth goal” accounts of assertion.

### 5.1 Objection Five: Graham’s Coherent Doubts

When Burge claims that an aspect of reason’s constitutive function is that of presenting truth, regardless of the reasoner’s self-interests, he claims to be introducing a “transpersonal” aspect of reason. You might object that Burge is equivocating between two different senses of ‘transpersonal’. Peter Graham develops this objection in his (2018).<sup>1</sup>

---

<sup>1</sup>Graham actually raises several objections to Burge’s overall project. It’s just that this specific objection, about a potential equivocation on the senses of ‘transpersonal’, is the objection that is specifically targeting aspects of Burge’s functional unity argument. The others have to do with whether Burge is in a position to establish that testimony is *ultima facie a priori* unconditionally reliable. Allegedly, that project fails because the reliability of testimony is conditional on the reliability of sense-perception, and Graham reads Burge as having conceded that the reliability of sense-perception, if it is reliable, isn’t *a priori*. For this discussion, see Graham (2018), pp. 101-106

Graham begins his objection by citing Burge's own remarks on the *apriority* of the transpersonal aspect of reason. In a footnote, Burge writes:

Although I think that my claim about this constitutive function of reason is *a priori*, I do not maintain that it is self-evident. It can and has been coherently questioned... But the claim has substantial initial plausibility, and I believe that this plausibility is deepened through reflection, including reflection on challenges to it. (1993, p. 475)

Leveraging this footnote, Graham's objection is simply to doubt both the initial plausibility of Burge's claim about the constitutive function of reason and he doubts that it survives challenge. I won't address the first doubt here because I have covered the initial plausibility of this claim; Graham's discussion of that doubt treads the same ground.<sup>2</sup> I'll instead focus on Graham's doubt that it survives challenge.

Graham's challenge starts with the concession that there are at least two aspects of reason, a theoretical aspect and a practical aspect<sup>3</sup>. The practical aspect is that of satisfying, i.e., making true, a person's plans and interests. The theoretical aspect is the transpersonal aspect mentioned two paragraphs ago. Having conceded that there's a theoretical aspect, Graham responds that there are two distinct claims that Burge could be making when he claims that there's a transpersonal aspect. On one sense, an aspect of reason's constitutive function is that of presenting truth to *the reasoner*, regardless of their own plans and interests. Let's call this the "merely-impersonal" aspect of reason. On another sense, an aspect of reason's constitutive function is that of presenting truth to *others*, regardless of *the reasoner's* own plans and interests. Let's call this the "truly-transpersonal" aspect of reason. We might, if we like, think of the "truly-transpersonal" aspect of reason as the merely-impersonal aspect of reason being extended to apply to

---

<sup>2</sup>See Graham (2018), pp. 107-108 .

<sup>3</sup>Graham (2018), p. 108

assertions as well.

Among other things, the distinction between the merely-impersonal and the truly-transpersonal aspects of reason helps Graham to coherently question the apriority of Burge's claims about the constitutive aspects of reason. You may concede that a person's reason(s) constitutively functions to present the truth to them while denying that it constitutively functions to present the truth to others. One might grant the apriority of the merely-impersonal aspect of reason while denying the truly-transpersonal aspect.

In support of granting the apriority of the merely-impersonal aspect while denying the apriority of the truly-transpersonal aspect, Graham offers a contrast between how belief and action relate to the will. He points out that action is a product of the will whereas belief is not. As he sees it, the will produces actions by forming and executing intentions. For Graham, this connects actions to the practical aspect of reason.<sup>4</sup> In contrast, beliefs are formed not by the formation and execution of an intention; rather they're rational reactions to the acquisition of evidence, to reasoning about the evidence, and the functioning of other belief-forming systems. For Graham, this connects beliefs to the merely-impersonal aspect of reason.<sup>5</sup>

Further, it seems right to say that people's assertions are actions. For it seems as though a person asserts a proposition upon forming and executing an intention to assert that proposition. But if assertions are actions, thinks Graham, then they fall under the practical aspect of reason rather than the theoretical aspect of reason (more specifically, the truly transpersonal aspect of reason).<sup>6</sup> That would mean that we lack a basis for saying that assertions also fall under the theoretical aspect of reason. That is, we don't seem to have a basis for extending the merely-impersonal aspect of reason into a truly-transpersonal aspect of reason.

---

<sup>4</sup>Graham (2018), p. 109

<sup>5</sup>ibid.

<sup>6</sup>Graham (2018), p. 109

Graham concedes that Burge could respond by claiming that assertion isn't like other sorts of acts. Unlike the acts of reaching for a glass of water, there's seems to be a constitutive connection between the act of asserting and the expression of belief. Put plainly, assertions seem to constitutively be expressions of belief. To Graham, this suggests the following line of argument:<sup>7</sup>

**BR1** Theoretical reason (i.e., belief) has an impersonal function of presenting truth

**BR2** One expression of theoretical reason is assertion

**BR3** So, reason functions in one mind to as to present truth to another mind through assertion.

Graham claims that if the transition from BR1 and BR2 to BR3 is successful, then Burge has a case for there being a truly-transpersonal aspect of reason.

Graham denies that the argument from (BR1)-(BR3) is persuasive. I reconstruct his complaint thus. For the argument to work, Burge needs to further defend a suppressed premise,

**BR4** If BR1 and BR2 are true, then BR3 is true

The addition of BR4 to the argument from BR1 to BR3 would result in a valid argument. What's at issue is whether it is sound. Graham doubts that it is. He writes:

We can still grant the impersonal function of reason, grant that assertion expresses states governed by impersonal reason, and so grant that assertions are a priori reliable signs of the normal operation of rationality (theoretical and practical), yet rationally doubt whether impersonal theoretical reason has the function of transpersonally presenting truth from one mind to another.

---

<sup>7</sup>Graham (2018), p. 110

Assertions are still actions, governed by practical reason.

Graham (2018), p. 110

And that concludes Graham's objection to the functional unity argument. Let's unpack what we can. Graham grants BR1 and BR2. He grants that BR1 and BR2 entail that assertions are *a priori* reliable signs of reason functioning normally. Nevertheless, he denies that a truly-transpersonal aspect figures into reason's functioning normally. Despite the conjunction of BR1 and BR2, assertions are still governed by the practical norms. Assertions, like all other actions, function to make true the reasoner's plans and desires. I.e., Burge hasn't established BR3.

I think Graham's objection to Burge's Functional Unity Argument is the most *prima facie* persuasive of the objections considered so far. I'd say Graham's objection is so *prima facie* persuasive because of the content of BR2. BR2 is a version of what we might call a belief-expression account of assertion. On a belief expression view of assertion, assertions are defeasible expressions of belief. This means that, absent indication to the contrary, a person's asserting some proposition indicates that they believe that proposition.<sup>8</sup>

For example, if I assert that that cat is hungry, absent indication to the contrary, you have *prima facie* evidence for believing that I believe that that cat is hungry. This defeasibility entails that that assertions do not necessarily co-exist with their associated beliefs. After all, it can be false that I believe that that cat is hungry, but true that I assert that that cat is hungry. So understood, the belief-expression view presupposes that the assertion and belief are distinct. This is an attractive result anyway because assertions are overt acts and beliefs are mental states.

The distinction between assertion and belief is a principled consequence of how assertions are things people do for reasons. This is just what Graham means when he says that assertion is governed by practical reason. A person can practically reason to

---

<sup>8</sup>For sophisticated views along these lines, see Bach (1979) and Grice (1957).

deceive people by asserting a lie. In doing so, they're precisely exploiting the *prima facie* defeasible link between what a person asserts and what they believe. Such a person might be doing this because they believe they plan to achieve something that is easier to accomplish if they lie. I submit that, given the belief-expression account of assertion, Graham has reason to deny BR4. Accordingly, Burge hasn't established BR3.

The ultimate heart of Graham's objection is that Burge needs to say something to explain how the truth of BR1 and BR2 is sufficient for the truth of BR3. Absent such a story, the doubts that Graham raises are overpowering. Everyone will agree that assertion, as a species of action, is subject to the practical aspects of reason. Burge needs to say more about assertion in order to motivate how it also falls under the jurisdiction of the impersonal aspects of reason.

## 5.2 Reply to Objection Five

BR2 expresses the belief-expression account of assertion. Such an account of assertion ultimately undermines the potential case for a truly-transpersonal aspect of reason. It does so because it does not indicate in any illuminating way how assertion might fall under the impersonal aspect of reason. Without such a story, Burge has not explained the case for a truly-transpersonal aspect of reason. Hence, the best response on Burge's behalf is to replace BR2 (and hence, BR4) with an account of assertion that does illuminate how assertion might fall under the jurisdiction of the impersonal aspect of reason.

Luckily for Burge, the belief-expression account of assertion isn't the only view on assertion. There are accounts of assertion that could explain a *prima facie* correlation between assertions and belief without taking it to be constitutive of assertion. One such account says presenting truth to others is the constitutive goal of assertion. Call this the "truth-goal" account of assertion.

I think the truth-goal account can be deployed by Burge to nonempirically explain how assertion, qua action, is governed by both a truly-transpersonal and a practical aspect of reason. The truth-goal account of assertion was introduced by Dummett in his (1959) and defended by Glanzberg in his (2003).<sup>9</sup> I propose to look at how Burge might, by assimilating Glanzberg's proposals, nonempirically explain how assertion functions to present truths. If he can do that, he would be in a position to give a compelling answer to Graham's doubts.

Here is Glanzberg's defense of the truth-goal account of assertion:<sup>10</sup>

**GD1** Contents determine satisfaction conditions

**GD2** Speech acts have *intrinsic purposes* determined by their constitutive practices

**GD3** If contents determine satisfaction conditions and speech acts have *intrinsic purposes* determined by their constitutive practices, then an intrinsic purpose of assertion is to present propositions as true.

**GD4** Therefore, the intrinsic purpose of assertion is to present propositions as true.

I'll explain each of these in order. GD1 is a common view of the representational contents of speech acts and propositional attitudes such as beliefs. Like I said in response to Shieber's objection; representational contents determine satisfaction conditions. In the case of beliefs and assertions, the satisfaction conditions of their contents are the conditions in which those contents are true. Hence, GD1 seems warranted.

---

<sup>9</sup>Glanzberg couches his defense of the truth-goal account as an intuition that there are no substantive "truth gaps". His idea is that a truth-goal account of assertion supports the thesis there is no meaningful sense in which a proposition is neither true nor false. I propose to set that aspect of Glanzberg's project aside. I will say nothing about whether there are any substantive "truth gaps", whatever that's supposed to mean. Instead, I will focus on the defense of the truth-goal account Glanzberg himself gives. This is not at odds with Glanzberg's work because he himself notes that the defense of the truth-goal account and his defense of the claim that there aren't meaningful "truth gaps" are separate.

<sup>10</sup>Glanzberg (2003), p.159

Burge can give a *prima facie a priori* gloss on the justification of GD1. Many mainstream accounts of content entail something like GD1. It doesn't matter which one we choose. One such view of representational contents is the unstructured possible worlds accounts. On this view, contents are equivalent to the sets of possible worlds that comprise their satisfaction conditions. Stalnaker (1999) is arguably an influential modern statement of the reductionist view. For this kind of proposal, GD1 is a trivial consequence.

There are also structured accounts of contents, according to which contents are distinct from the sets of possible worlds that comprise their satisfaction conditions. You might be attracted to a structured account of content because you think that there can be two or more distinct contents that nevertheless yield the same satisfaction conditions. One historical statement of this view is found in Frege's (1892/1997).

These days, there at least two competing structured accounts. There are neo-Russellian accounts, according to which contents are sequences composed of particulars. Modern defenses of this view include David Kaplan's (1989) and Nathan Salmon's (1986). There are also neo-Fregean accounts, according to which contents are instead composed of sequences of "modes of presentation", i.e., ways of representing particulars. Defenses of the neo-Fregean accounts include David Kaplan's earlier (1968), as well as Burge's own responses in his (1977) and (2009a).

For structured accounts of propositions, GD1 is still a substantive consequence. Let's consider how this might be handled in the context of a neo-Fregean account. On a neo-Fregean account, there are at least two important facts about content.

The first important thing about contents on a neo-Fregean account is that contents determine satisfaction conditions. Here is a provisional gloss on a content's satisfaction conditions. In the case of the contents of beliefs and assertions, the satisfaction conditions for such contents are the circumstances in which the relevant contents are true.



For example, suppose that Kat believes Mittens the cat is asleep on their bed. That belief is true in all and only those circumstances in which Mittens the cat is asleep on their bed. Those circumstances comprise the satisfaction conditions for the content that Kat believes. Alternatively, we can think of satisfaction conditions as rules associating circumstances with truth-values.<sup>11</sup> Consider again the example where Kat believes that Mittens the cat is asleep on their bed. The veridicality conditions for that belief's content is a rule pairs all and only those circumstances in which Mittens the cat is asleep on their bed with the veridicality value for truth.

The second important thing about contents on a neo-Fregean account is that the composition of contents explains the composition of satisfaction conditions. This is achieved by way of the Fregean account of predication. On any Fregean account of contents, complex contents are composed by “function” application.<sup>12</sup> The intuitive idea is that contents with satisfaction conditions are composed by applying simpler contents to one another using certain rule governed operations. The satisfaction conditions for a content bearing state are determined in a computable way by these rule governed operations. Hence, on a neo-Fregean, even though contents are not equivalent to their satisfaction conditions, they play an important explanatory role in determining those satisfaction conditions. Hence, according to a neo-Fregean account GD1 is true as a substantive, explanatory consequence.

Regardless of which proposal you pick, these proposals are all *a priori* proposals about logic and semantics. They're motivated, at least in the first place, by *a priori* reflection on how meanings contribute to logical deduction and mental representation. Since Burge himself already gives an *nonempirical* defenses of his preferred Neo-Fregean account of

---

<sup>11</sup>The account of satisfaction conditions as rules associating worlds with veridicality is due to Alonzo Church (1940).

<sup>12</sup>This is one of the main innovations of Church (1940)'s function application account of content. For a detailed review of how Church's function application account informs semantic theory in modern linguistics, especially in neo-Fregean semantics, see Heim and Kratzer (1998).

content, we can safely say that GD1 is *prima facie a priori*.

GD2 says that speech acts have *intrinsic purposes* determined by their constitutive practices. This is a view of assertions, and general language use, that enjoys wide purchase in the philosophizing about assertion. Let's call this broader view of language use the "practice view". According to the practice view, assertion and many other central speech acts are intrinsically purposive because of their constitutive linguistic practices. Linguistic practices, in turn, are constituted in part by their rules and their overarching intrinsic purposes. When Glanzberg says that some act or activity has an intrinsic purpose, he means means that they have essential functions independent of the speaker's own-interests; these functions instead come from the nature of the act itself.<sup>13</sup> The functions a particular action has in virtue of the agent's interests are "extrinsic". So construed, the intrinsic functions, which an action has in virtue of the type of act it is, are independent of the speaker's own self interests. Intrinsic purposes are, in that sense, impersonal.

On the practice view deployed in GD2, an act's intrinsic purpose is determined by its constitutive practice, i.e., the practice relative to which the act is individuated. A practice is a type of activity that's individuated in some part by its rules and by its intrinsic purpose.<sup>14</sup> Games are often offered as the paradigmatic species of practices.<sup>15</sup> To make things simpler, I'm going to focus on a specific sub-species of practices: games. I will do so because games are great exemplars of the explanatory roles of a practice's intrinsic purposes/goals and rules.

Games are a paradigmatic type of practice. Games are practices where the general, intrinsic purpose is that of "winning". Not all practices are games in this sense. Winning

---

<sup>13</sup>Glanzberg (2003), p. 160.

<sup>14</sup>Rawls (1955), p. 24; Glanzberg (2003), p. 161.

<sup>15</sup>See Dummett (1959), Brandom (1994), and Williamson (2000) for influential examples of this trend.

occurs when a game concludes with one party gaining victory over the others.<sup>16</sup> A game's rules specify its victory conditions. For example, the rules of the game of NBA regulation basketball specify that game's victory conditions. In NBA basketball, winning consists in one team's having produced more points than the other at the expiration of play time.

The rules also explain why act types specific to a practice necessarily have intrinsic purposes. The general act of tossing a ball need not have an any purpose beyond the extrinsic purpose the agent imposes upon it. That's because the act of tossing a ball isn't specific to a practice. In contrast, the jump shot has the intrinsic purpose of promoting the shooter's team score. The jump shot has this intrinsic purpose because it is individuated by its place in the practice of basketball. A jump shot has this purpose, even if the player performing it merely intends to draw a shooting foul, with no expectation of successfully increasing the score with that jump shot. A player may do this, for example, if they're hoping to sacrifice their attempted jump shot for a more tedious, but reliable, series of foul shots. Foul shots, too, have the intrinsic purpose of increasing the score of the agent's team. The rules of the game elucidate how else jump shots and foul shots differ.

Further, the intrinsic purpose of a game explains what it is for a person to intentionally play that game.<sup>17</sup> If a person can play intentionally, they can play to win. If they can play to win, they're sensitive to the game's intrinsic purpose of winning. For example, a person who isn't sensitive to the goal of winning in NBA regulation basketball isn't able to intentionally participate in NBA regulation basketball, because they can't play to win. I invoke the notion of sensitivity where Glanzberg invokes the notion of understanding. I think sensitivity is the more accurate notion because it is less psychologically demanding.

---

<sup>16</sup>I think this is so, even in cooperative and solitary games. For example, in solitaire, the other "party" is procedurally generated. This procedurally generated party is the array of card stacks generated by shuffling and drawing cards in a pre-arranged pattern.

<sup>17</sup>Glanzberg (2003), pp. 161-162.

Sensitivity to the goal of winning requires that a person alter their behavior to maintain pursuit of the goal, even if they are ignorant or mistaken about some aspects of it. For example, many players learn games by playing; they strive to win even if they don't fully understand all aspects of what winning requires. Understanding the goal of winning requires more than this. Understanding requires that a person be able to veridically describe substantial aspects of what winning requires and how to achieve it in varying conditions.

Now consider a person who is not, in anyway, sensitive to the intrinsic goal of NBA regulation basketball. This would be a person who is not, in anyway, sensitive to the goal of producing more points before the end of competitive play. Intuitively, this is a person who is not able to participate in the game of NBA regulation basketball. Such a person would not be sensitive to the pressures that constitute competitive play in NBA regulation basketball.

Glanzberg follows several other philosophers in advocating a practice view of language use.<sup>18</sup> Language use appears to be an intrinsically purposeful activity constituted by rules.<sup>19</sup> For example, Grice's basis for proposing his celebrated Cooperative Principle seems to be an *a priori* intuition the overarching intrinsic purpose of all forms of conversation is the coordination of multiple agent's behavior in the service of a joint endeavor.<sup>20</sup>

Others, loosely inspired by Grice's account of general conversation, advocate narrower intrinsic purposes for narrower species of conversation. Two prominent examples of this are the practices of inquiry and rational dialectic.<sup>21</sup> In regards to inquiry, Robert

---

<sup>18</sup>Glanzberg (2003), p. 162. See also Stalnaker (1978), Lewis (1979), Brandom (1994), Williamson (2000), Roberts (2012), and Rescorla (2009a).

<sup>19</sup>For example, see Grice (1989), Lewis (1979), and Brandom (1994).

<sup>20</sup>Grice (1975), p. 45

<sup>21</sup>I think there are fair questions as to whether inquiry and rational dialectic comprise genuinely distinct language games. I'm inclined to say that although they're superficially distinct, insofar as their intrinsic purposes seem to be different, they share deeper structural elements. For a first pass at what

Stalnaker and Craige Roberts propose idealized models on which inquiry is that language game where the intrinsic purpose or goal consists in the settling of what questions are under discussion. For example, Stalnaker writes,

“The motivation for representing the speaker’s presuppositions in terms of a set of possible words... is that this representation is appropriate to a description of the conversational process in terms of its essential purposes. To engage in conversation is, essentially, to distinguish among alternative possible ways that things may be”.

Stalnaker (1999) p. 85

I think this is a constitutive claim about inquiry because claims about essence are claims about what’s constitutive of what. In her work building on Stalnaker’s initial insights, Roberts also reads Stalnaker as explicitly claiming that inquiry is intrinsically a matter of settling questions about the world.<sup>22</sup> At the very least, Roberts seemingly reads Stalnaker as I do.

In regards to rational dialectic, Robert Brandom proposes an account in which rational dialectic’s intrinsic purpose consists in the settling of the “deontic score”, i.e., the distribution of commitments and responsibilities among interlocutors. You might also remember our earlier discussion of Michael Rescorla’s No Asymmetries Argument.<sup>23</sup> Over the course of that argument, Rescorla develops the view rational dialectic’s intrinsic purpose consists in players settling on rational agreement while advancing positive reasons and avoiding decisive defeat.<sup>24</sup>

Just as conversation and the conversational activities of reasoned dialectic are appear

---

those deeper structural similarities might be, see Walton and Krabbe (1995)

<sup>22</sup>Roberts (2012) p. 5

<sup>23</sup>Rescorla (2007)

<sup>24</sup>For other works elaborating upon other aspects of rational dialectic, see MacFarlane (2003) and Kolbel (2011). Rescorla’s discussion develops insights from the deeply underrated Walton and Krabbe (1995)

to be constitutively or intrinsically purposeful, so too does assertion appear to be constitutively intrinsically purposeful. Assertion appears to be so purposefully because such purposefulness explains what it is for people to form intentions to perform speech acts such as assertions.<sup>25</sup> A person, for example, cannot intend to lie without some sensitivity to how assertion intrinsically functions to convey the information that the expressed propositional content's associated satisfaction conditions are satisfied.<sup>26</sup> A person who is insensitive to the intrinsic purposes of assertion, and how those intrinsic purposes fit into broader conversational activities, is someone who is unable to participate in any of the activities of asserting and conversing. This would be akin to a person who has at their disposal all the sentences of English but is completely insensitive to the purposes to which those sentences are put. Such a person has no entry point into how those sentences are used or what those sentences really mean. After all, sentence meanings, like sentences themselves, are abstracta that partly explain how sentences are used.<sup>27</sup> But those uses are essentially tied to practices, i.e., intrinsically purposeful activities. Since the person in question is not sensitive to those intrinsic purposes, they would not be able to do anything with all those sentences of English. Sentences of English would be, for him, meaningless gibberish.

I think that Burge can also provide a *prima facie a priori* gloss on the justification of GD2. I grant that this might not be obvious. This might not be obvious because a linguist might deploy some variant of the practice view to form empirical theories about specific patterns of linguistic behavior. And the explanatory success of such an empirical theory might provide *prima facie a posteriori* justification for the practice view.

Nevertheless, I deny that the practice view, and its proponents, are exclusively mo-

---

<sup>25</sup>Glanzberg (2003), pp. 162-163.

<sup>26</sup>*ibid.*, p. 163

<sup>27</sup>I do not mean to endorse the view that there is nothing more to the meaning of a sentence than its use. For such a proposal, see Wittgenstein (1953).

tivated, in the first place, by specific citations of specific empirical studies of specific languages. Instead, the practice view's proponents often support their view with *a priori* reflection on thought experiments and our intuitions.<sup>28</sup> These reflections are *prima facie a priori* because they're not tied to reflection on the specific details of specific languages. Instead, they're tied to our pre-theoretical understanding of what it is to be a language user. The practice view, at least in the hands of philosophers such as Brandom, Dummett, Glanzberg, Grice, Lewis, and Stalnaker, is supported by such *prima facie a priori* reflections. This methodology isn't substantively different from the sorts of *a priori* reflections Burge sometimes uses to support his views about the nature of contents.<sup>29</sup> In light of those methodological reflections, I think that Burge can endorse GD2 on *prima facie a priori* grounds.

GD3 says that if GD1 and GD2 imply that an intrinsic purpose of assertion is to present propositions as true. Elaborating further, Glanzberg claims:

...we may conclude that the intrinsic purpose of assertion is somehow given in terms of the proposition—the content—expressed by an assertion. How? We have already observed that the purpose is not adequately described as simply determining a collection of truth conditions. This does not tell us what the speaker is attempting to do with these conditions... we must add that the purpose is to say that the truth conditions given by the proposition expressed obtain. In conveying the information that the truth conditions of 'The cat is on the mat' obtain, the assertion of this sentence describes the world as being some way—namely, one in which the cat is on the mat. The obtaining

---

<sup>28</sup>Wittgenstein (1953) for an influential, though heavily stylized, instance of this kind of methodology. I grant that Wittgenstein would not describe his project this way. Still, the abstract, stylized passages where Wittgenstein introduces the notion of a language game strike me as *prima facie a priori* reflections on arbitrarily introduced thought experiments. He does not seem to be reflecting on specific real world studies.

<sup>29</sup>(1979)

of truth conditions thus provides the intrinsic purpose of assertion.

Glanzberg (2003), pp. 163–164.

I reconstruct Glanzberg's underlying reasoning as follows. First, a person's assertion essentially involves them using a *declarative sentence* in context. This is plausibly nonempirical. When we think about what it is for a person to assert a proposition in a conversational context, it involves, among other things, the person using a declarative sentence in that context.

Second, a person's using a declarative sentence in context expresses a proposition.<sup>30</sup> This is plausibly nonempirical. When we think about what it is to assert a proposition in a conversational context, the proposition asserted is what's expressed by using a declarative sentence on that context.

Third, the purpose of conveying the information that the satisfaction conditions conveyed are satisfied may coincide with the speaker's designs, but it is intrinsic to asserting itself. This is also plausibly nonempirical. Just as a basketball player's plans in performing a jump shot involve a know-how sensitivity to the intrinsic purpose of a jump shot, so too does a speaker's plans in asserting a proposition involve a know-how sensitivity to the intrinsic purpose of asserting.<sup>31</sup> Here, I am helping myself to the plausible assumption that presenting a proposition as true consists in purportedly conveying the information that the satisfaction conditions conveyed (i.e., determined by that proposition), are satisfied.

Fourth, and finally, the intrinsic purpose of assertion isn't exhaustively describe merely in terms of conveying satisfaction conditions. Rather, the intrinsic purpose of assertion is characterized in terms of conveying that those satisfaction conditions are

---

<sup>30</sup>Setting aside, of course, cases like reference failure.

<sup>31</sup>Again, I depart with Glanzberg on how to describe the possession conditions for the ability to assert. Understanding strikes me as overly demanding. Most speakers haven't even thought about the intrinsic purpose of assertion, and still they perform them. Still, their performances display a sensitivity to how assertions intrinsically function to convey information the that *p*.



indeed satisfied. Here, there are at least two optional, though admittedly more controversial ways, that a supporter of Glanzberg might try further support the fourth and final point.

One way Glanzberg might support the fourth and final point is by appeal to an old Fregean idea. The old Fregean idea is that two speech acts can express one and the same proposition, i.e., convey one and the same set of satisfaction conditions, and nevertheless differ in their force. For Frege, the original contrast is between assertion and commands. I may, for example, assert that the trash is outside. Alternatively, I may command that the trash be outside. Both of these express the same propositional content, which is true in all and only those situations where the trash is outside. Nevertheless, the force with which these satisfaction conditions are presented (conveyed) is different. In the case of assertion, I am explicitly presenting as true the proposition that the trash is outside. In the case of command, I am explicitly presenting that proposition as one to be made true by the addressee. Accordingly, the mere conveyance (expression) of a proposition and its satisfaction conditions is not enough. We must also specify that assertion's intrinsic purpose is to present propositions as true (i.e., convey that the satisfaction conditions expressed are satisfied).

Another way that Glanzberg might support the fourth and final point is by appealing to the broader practice within which assertion occurs. As previous discussions suggest, Glanzberg has options here. He might appeal to the Stalnaker-Roberts account on which assertion is essentially situated within a practice of inquiry.<sup>32</sup> Alternatively he might appeal to Brandom and MacFarlane's accounts on which assertion is essentially situated within a practice of rational persuasion.<sup>33</sup> Alternatively, Glanzberg might follow Walton and Krabbe and Rescorla's accounts on which assertion is individuated by its place in a

---

<sup>32</sup>Stalnaker (1999) and Roberts (2012)

<sup>33</sup>c.f. Brandom (1994) and MacFarlane (2003).

family of practices, reasoned discourse, which subsumes both inquiry and persuasion.<sup>34</sup> These three accounts all situate assertion as a move in a practice of giving reasons in support of the truth of some position. In all of these accounts, assertion's role is that of explicitly advancing some proposition as true. This is how assertion, in the practice of reasoned inquiry, contributes to the settling of the question under discussion. In such a setting, a person's assertions present propositions as true because the truth of those assertions support settling the question under discussion in favor of some answers over others. This is also how assertion, in the practice of reasoned persuasion, contributes to persuading participants of one position versus another. In such a setting, a person's assertions present propositions as true because the truth of those propositions constitute reasons to agree with some participants' positions versus others. Advancing some proposition as true just is explicitly presenting it as true.

An advantage of Glanzberg's enriching his account of assertion in the ways suggested by Brandom, MacFarlane, Walton and Krabbe, and Rescorla, is that it allows Glanzberg to say more about what it is to explicitly present a proposition as true. On these dialectical, or reasoned discourse, accounts of assertion, we get the following account of what it is to explicitly present a proposition as true. A person explicitly presenting a proposition as true is explained, at least in part, by their undertaking a commitment to the truth of that proposition. As Walton and Krabbe explain, a commitment to the truth of a proposition is a specific sort of *action commitment*.<sup>35</sup> An action commitment is a relationship between a subject and an object. The subject of an action commitment is the person or collective who bears the commitment. The object of an action commitment is the course of action that the subject is committed or bound to do.

We might think of person's action commitments as essentially associated with imper-

---

<sup>34</sup>C.f. Walton and Krabbe (1995) and Rescorla (2007).

<sup>35</sup>Walton and Krabbe (1995) p. 15

atives or commands addressed to that person. A person satisfies a command addressed to them insofar as they take steps to make true the content of that command. A person “lives up to an action commitment”, i.e., satisfies the commitment, to the extent that execute, i.e., make true, the corresponding command addressed to them.<sup>36</sup>

Armed with the notion of an action commitment, Walton and Krabbe can then say what propositional commitments consist in. A person’s propositional commitments are then action commitments that enjoin the commitment bearer to take certain practice specific actions that center upon a specific proposition.<sup>37</sup> In the case of assertion, a person who asserts a proposition, i.e., who explicitly presents a proposition as true, is a person who purports to explicitly undertake a commitment to either explicitly defend the proposition in question from decisive challenges and counterexamples, or else explicitly retract their assertion of that proposition.<sup>38</sup>

At any rate, these considerations for enriching Glanzberg’s argument are merely optional additions. After all, I am mainly interested in Glanzberg’s argument as a way of independently establishing that assertion, like belief, is subject to the impersonal aspect of reason. If the considerations that support the claim that truth is the intrinsic purpose of assertion also support that knowledge is an intrinsic standard for asserting, I would be satisfied. For knowledge is an achievement of the impersonal aspect of reason, maybe even, the central achievement. All that matters for my purposes is giving reasons to believe that assertion falls under the jurisdiction of the impersonal aspect of reason. The most straightforward way for this to be plausible is for explicitly presenting the truth being the intrinsic goal of assertion. This could easily also be achieved by a knowledge rule for assertion.

Let’s summarize what has been established so far. GD1 established that contents

---

<sup>36</sup>Walton and Krabbe (1995) pp. 16-21

<sup>37</sup>Walton and Krabbe (1995) pp. 23

<sup>38</sup>Walton and Krabbe (1995) p. 12, p. 31

convey satisfaction conditions because they determine satisfaction conditions. GD2 established that assertions are intrinsically purposeful/functional because they're specific to some language game, perhaps the language game that encompasses inquiry and rational dialectic. GD3 builds on GD1 and GD2 to establish that the intrinsic purpose of assertions is to convey the information that the satisfaction conditions conveyed in the utterance obtain. From there, it seems *prima facie a priori* that the intrinsic purpose of assertion is to present truth. A proposition's being true seemingly consists in its satisfaction conditions obtaining. Since GD1, GD2, and GD3 seem true, we can conclude that since GD4 also seems true. Accordingly, the intrinsic purpose of asserting is to present truth.

### 5.3 Rejoinder to the Reply to Objection Five

I have offered (my take on) Glanzberg's argument for the claim that assertion is governed by an intrinsic purpose of presenting truth. I think that armed with this argument, Burge can amend the argument below for the claim that both belief and assertion fall under the jurisdiction of the impersonal aspect of reason. This argument was initially offered by Graham on Burge's behalf:<sup>39</sup>

**BR1** Theoretical reason (i.e., belief) has an impersonal aspect of presenting truth

**BR2** One expression of theoretical reason is assertion

**BR4** If theoretical reason (i.e., belief) has an impersonal aspect of presenting truth, and if one expression of theoretical reason is assertion, then reason functions in one mind so as to present truth to another mind through assertion.

---

<sup>39</sup>Graham (2018), p. 110

**BR3** So, reason functions in one mind to as to present truth to another mind through assertion.

And as you recall, this argument failed because it is plausible to doubt the truth of BR2 and BR4. After all, Burge has said nothing to cement BR2 or BR4 as plausible. An opponent of Burge can grant BR1 and BR2 and still doubt BR4.

I think that, armed with Glanzberg's argument, Burge can amend the argument above as follows:

**Glanzberg-Addendum to The Functional Unity Argument**

**BR1** Belief has an impersonal aspect of presenting truth

**GR4** Assertion also has an impersonal aspect of presenting truth

**GR5** If belief and assertion each have an impersonal aspect of presenting truth, than reason functions in one mind so as to present truth to another mind through assertion.

**BR3** Reason functions in one mind so as to present truth to another mind through assertion.

This is the argument that Burge needs to support the functional unity argument's assumption that assertion is also subject to the impersonal aspect of reason.

In light of this revision, I think there's still one more move that someone enamored with Graham's doubt objection can deploy. A proponent of the doubt objection might still, nevertheless, insist that the premises of this argument can be coherently doubted. I think there are a couple of ways of pitching this objection. One way is to pitch the objection is to doubt that BR1 and GR4 cite one and the same impersonal aspect of presenting truth. However, I think the better way is to simply pitch the objection as

doubting the truth of GR5. I will consider how to raise the doubt as described in this latter way.

Let's consider how someone might doubt the truth of GR5. That is, let's consider how someone might concede that belief and assertion each have an impersonal aspect of presenting truth while doubting that reason functions in one mind so as to present truth to another mind through assertion. I think that here's a cogent way to present the doubt. When I considered Glanzberg's argument for the claim that an intrinsic purpose of assertion is that of presenting truth, we were talking about how the practices constitutive of assertion determined how assertion intrinsically functions to present truth. Now, we might think that this intrinsic function of presenting truth doesn't unify with the corresponding function for belief. That is, we might grant that belief and assertion each have an impersonal aspect of reason, but doubt that both of these impersonal aspects must unify with the practical aspect of reason. It might be that the impersonal aspect of reason, as applied to belief, and the practical aspect of reason, as applied to action, have to unify. And it might also be that the impersonal aspect of reason, as applied to assertion, is exempt from this unity requirement. If so, then GR5 is not true.

Let's elaborate further on the plausibility of this doubt. At the heart of the doubt is a straightforward observation about Glanzberg's argument regarding the intrinsic purpose of assertion. That observation is the observation that Glanzberg's argument has to do with assertion as a facet of a practice. But practices are essentially complex act types. They're not obviously facets of agency. If they were, then the intrinsic purpose of assertion must unify with the other aspects of reason in a person's action. But there's nothing incoherent about positing a solitary rational agent who has no need, capability, or opportunity to engage in discourse with another agent. Such an agent would be one in a solipsistic situation. We might even imagine this solipsistic reasoner as able to engage in genuine action to satisfy their own needs.

I think that such rational agents are indeed possible. At the very least, I am happy to concede, at least for the sake of argument, that conceiving of such rational agents doesn't seem to me to raise any apparent contradictions. Thus, such solipsistic rational agents are at least logically possible.

Such a logically possible solipsistic rational agent might be able have conceptual access to a practice of discourse. Perhaps this agent, in a moment of self-consciousness, envisioned the possibility of others like themselves, and like Ariel of the Disney's take on the little mermaid, wondered what it would be like to be part of that world. In envisioning this possibility, this agent envisioned various rational agents engaged in our practices of conversation and reasoned discourse. Nevertheless, because this agent is a solipsistic one, the practices of conversation and reasoned discourse are divorced from its agency. They are, at best, external environmental conditions that have no bearing on that agent's rational nature. This is why it seems plausible to me that such a rational agent is not required to unify the intrinsic purposes of assertion with the other aspects of reason.

Here is another consideration in support of the doubt in question. The account of assertion that emerges in support of GD4 is an account of assertion that does not obviously say anything about what to believe. You might think that a person who, under the grip of a conspiracy theory, disbelieves the truth, can nevertheless be a good asserter. Such a person will assert what is in fact the truth not because they believe that it is the truth but because they believe that the powers that be will punish him or her if he or she reveals himself as believing otherwise. Essentially, this person asserts the truth, but for the wrong reasons. They assert the truth not because they believe that it is the truth, but because they believe that it is an instrumentally advantageous falsehood. Nevertheless, this would be a person who seems to be asserting well, at least with respect to the intrinsic purpose of assertion.

A similar example would be the case of the famed case of the Creationist Teacher that comes up in other areas of the epistemology of testimony. In the creationist teacher scenario, we're to imagine a competent biology teacher who can apply and teach the main tenets of mainstream evolutionary theory. At the same time, because this teacher is a creationist, they nevertheless disbelieve the main tenets of mainstream evolutionary theory. On a field trip with their students, the creationist teacher uncovers a new fossil and correctly asserts to their students that what they have found is a fossil.

In the hands of its creators, this creationist teacher case is leveraged as part of an argument against the view that testimony is a purely preservative channel of warrant.<sup>40</sup> I am ambivalent on the status of that particular debate.<sup>41</sup> I am more interested in how the creationist teacher case make plausible doubts to GD5 in the Glanzberg-inspired addendum to the functional unity argument. The creationist teacher case seems to suggest that a person can satisfy the intrinsic purpose of assertion without being unified. Insofar as we would say that this person was doing their job well in asserting the truth, we can say that this is a case where the person has satisfied the intrinsic purpose of assertion while falling short of the impersonal aspect of reason, as applied to belief.

Both the case of the solipsistic agent and the creationist teacher seem to make plausible the doubts concerning GD5. Even if both belief and assertion are beholden to an impersonal aspect of reason, it doesn't seem to follow, at least all by itself, that a person is required to unify both those aspects with the aspect of practical reason.

---

<sup>40</sup>c.f. Lackey (1999); Graham (2000)

<sup>41</sup>For a promising line of response, see Burge (2013) pp. 254-264



## 5.4 Response to the Rejoinder to the Reply to Objection Five

The rejoinder to the reply to objection Five says that GD5 is susceptible to doubt. That is, a person grant that belief and assertion both have an impersonal aspect of reason and yet deny that those aspects must both be unified with the practical aspect of reason. At the heart of the doubt is this idea that there's a distance between rational agency and the practice of asserting. This distance isolates the impersonal aspect of reason, as applied to asserting, from the unity requirement on ideal rationality.

Let's deliver a final response to this rejoinder. In delivering this final response, my goal is not to evangelize. I do not aspire to convert opponents of my view to my side. Instead, I simply want to provide the materials to resist the doubt raised by the rejoinder on their behalf.

I start by helping myself to the claims that both belief and assertion have an impersonal aspect of reason. By this, I mean that presenting truth is an intrinsic purpose of both belief and assertion. From here, I think we have two complementary options for establishing the further claim that reason functions in one mind as to present the truth to another mind.

The first of these complementary options goes like this. I concede that cases like the solipsistic rational agent and the creationist teacher are logically possible. Nevertheless, I think that assertion's broader individuating practice provides practitioners an in principle default reason to assert sincerely. And I think that to assert sincerely is simply to assert what you believe. Since, on the current assumption, both assertion and belief are governed by an impersonal aspect of presenting truth, it then follows that the practice of assertion generates default rational pressure for a practitioner to be unified. I grant that the reference to "principle default reason" is the crucial controversial point, so let

me start by saying why I think that reference is appropriate.

My reason for thinking that assertion's broader individuating practice provides practitioners an in principle default reason to assert sincerely is an argument that I've cited earlier in this essay. Earlier, when I gave some reasons for thinking that a unified agent is a rational one, I brought up a battery of arguments that suggested that a practically rational agent who is also impersonally rational enjoys systematic advantages over practically rational agents are not, in some relevant way, impersonally rational. Among these arguments was Rescorla (2007)'s No-Asymmetries Argument. The No-Asymmetries Argument was Rescorla's attempt at explaining how the idealized conditions of rational discourse generates in-principle defeasible default rational pressure towards asserting sincerely rather than not.

Here's a brief refresher on Rescorla's No-Asymmetries Argument. Rescorla's No-Asymmetries Argument is offered within the context of a dialectical account of assertion. According to a dialectical account of assertion, assertion is individuated, at least in part, by its place in the "game of giving and asking for reasons".<sup>42</sup> The game of giving and asking for reasons is not so much a single, monolithic type of game or practice as it is a family of games or practices. The paradigmatic examples of games within this family this family include the rational persuasion and rational inquiry games.<sup>43</sup> The rational persuasion game is a game where participants attempt to persuade each other of their idiosyncratic views. The rational inquiry game is a game where participants attempt to settle questions of fact together. Intuitively, a single instance of real life discourse may instantiate both games.

Despite their differences, both the rational persuasion and the rational inquiry games are individuated in part by having the same set of constitutive goals.<sup>44</sup> The overarching

---

<sup>42</sup>C.f., Brandom (1994), MacFarlane (2003)

<sup>43</sup>Walton and Krabbe 1995

<sup>44</sup>Rescorla (2007)

constitutive goal is the achieving rational *rapprochement*, i.e., reasoned agreement. In the rational persuasion game, rational *rapprochement* is constituted by reasoned agreement as to whose idiosyncratic position is correct. In the inquiry game, rational *rapprochement* is constituted by reasoned agreement as to the correct answer to the question under discussion.

The constitutive goal of rational *rapprochement* is supplemented by two ancillary constitutive goals. One of these ancillary constitutive goals is the goal of making positive moves towards *rapprochement*.<sup>45</sup> This ancillary constitutive goal determines as inappropriate such stonewalling strategies as “merely asking questions”. After all, a person who stonewalls is a person who insists on partaking in a game of giving and asking for reasons and then stonewalls and then does anything other than giving reasons.

The second of these ancillary constitutive goals is the goal of avoiding decisive objections.<sup>46</sup> If you think about how this connects to achieving rational *rapprochement*, the goal of avoiding decisive objections makes sense. You cannot have reasoned agreement on something if the alleged reasons underwriting that agreement have been exposed to decisive objection. This is because the decisive objections to alleged reasons drain those alleged reasons of whatever rational force they may have. Let’s call the overarching constitutive goal of rational *rapprochement*, together with the two supplementary constitutive goals, the tripartite constitutive goals of reasoned dialectic. The tripartite constitutive goals of reasoned dialectic jointly individuate core cases of games of giving and asking for reasons.

In addition to the constitutive goals that unify games of giving and asking for reasons, there are constitutive rules that position assertion as a basic move in such games. These are the “Defense/Retraction” rules.<sup>47</sup> A simplified version of these rules might be stated

---

<sup>45</sup>Rescorla (2007)

<sup>46</sup>Rescorla (2007)

<sup>47</sup>c.f. Brandom (1994), MacFarlane (2003), Rescorla (2007), and Walton and Krabbe (1995)

as follows: A person must assert a proposition only if they undertake the commitment to either defend that proposition from challenge and counterexample or else to issue a retraction. Accordingly, suppose that Jane asserts that John's store is closed. In doing so, she must undertake a commitment to the following complex policy: either she defends the truth of the proposition that John's store is closed or else she must issue a retraction of her testimony. If, as the discourse grows, and as challenges to her assertion grow, Jane does not either defeat those challenges or admit a retraction of her testimony, then she will be playing defectively. If these defects are too great, it might not be plausible to say that Jane is even partaking in a game of giving and asking for reasons. Jane can take either route of defense or retraction to discharge her commitment. However, if Jane issues too many retractions, she falls in danger of stonewalling.<sup>48</sup>

The above example illustrates not only how a Defense/Retraction rule works, but also how it interacts with the constitutive goals of assertion to generate rational pressure towards some patterns of assertion over others. For example, there's pressure for good faith participants of a game of giving and asking for reasons to defend more assertions than they retract. That's because if they issue too many retractions, they fall short of the constitutive goals of the game and fall short of victory. And on the assumption that they're good faith participants of a game of giving and asking for reasons, they're playing to win.

So, on a dialectical account of assertion, assertion is individuated "by its place" in the games of giving and asking for reasons. Games of giving and asking for reasons are individuated by the centrality of three goals, which jointly describe what winning the game amounts to: the goal of achieving rational rapprochement, the goal of making contributions towards rational rapprochement, and the goal of avoiding decisive objections. Assertions get their place in such games by way of their individuating defense/retraction

---

<sup>48</sup>Rescorla (2007)

rules. Finally, these defense/retraction rules incentivize good-faith players to pursue certain dialectical strategies over others.

Keeping in mind how the defense/retraction rules, in concert with the tripartite goals of reasoned dialectic, incentivize good-faith players to pursue certain dialectical strategies over others, I turn to give my gloss on Rescorla's explanation for why it seems as though a person ought to assert only what they sincerely know. In everyday, ordinary life realizations of games of giving and asking for reasons, the players are rarely, if ever, on even footing with one another. Some players are more clever than others while other players know more about the relevant topics than others. In some contingencies, there will invariably be players who are able to assert lies probably get away with it.

Now contrast ordinary real life contingencies with the ideal conditions described in the game of giving and asking for reasons. In such conditions, players know nothing about each other. They do not know who among them is better at argumentation. They also do not know who among them is the superior expert on the subject matter at hand. Accordingly, assuming they're good faith players, i.e., players who participate with the goal of winning, they need to pick strategies that maximize their chances of winning while providing them a degree of protection from failure. In such idealized conditions, players are rationally entitled to presume that they enjoy no epistemic or cognitive advantages over other players. Further, in such idealized conditions, if players are rationally entitled to presume that they enjoy no epistemic or cognitive advantage over other players, then the ideally rational default strategy is for players to sincerely assert from their stock of warranted beliefs. And a person's warranted beliefs are the ones that they properly believe to be true.

Asserting sincerely from your stock of knowledge and warrant is the ideally rational default strategy in the idealized conditions of a game of giving and asking for reasons because it is the only strategy that that counts as advancing reasons for a position, i.e.,

avoids stonewalling, while at the same time maximizing your ratio of successful defenses to issued retractions. Here is why sincerely asserting from your stock of knowledge and warrant maximizes your ration of successful defenses to issued retractions. Every time you assert, you undertake a commitment to defend or retract the resulting testimony. But the only way to defend your testimony is to assert further propositions. Asserting further propositions incurs additional commitments to defend or retract additional testimony.

Now, suppose that in the idealized conditions of a game of giving and asking for reasons, you assert lies rather than from your stock of warrant. Because lies are disconnected from the truth, a full defense of a lie will invariably require further lies. That's because you need to tell an additional lie to paper over the original lie's disconnection from the truth. This means that a full defense of a lie generates a commitment to defend further lies. Each of these additional lies increases your attack surface in the sense that they increase the occasions in which you will be unable to defend and be forced to issue a retraction. And since those lies will be holding up other instances of testimony, you will be forced to retract that testimony as well.

Suppose in contrast that, in the idealized conditions of a game of giving and asking for reasons, you assert the truth from your stock of warrant rather than telling lies. When you do this, you're still vulnerable to challenges. And to answer those challenges, you still have to offer further assertions and thus incur additional commitments to defend or retract the testimony that results. However, unlike in the case where you try to give a full defense of a lie, you are asserting the truth from your stock of warrant. Your answers to challenges to your testimony's truth is to simply reproduce your access to that truth.

Rescorla's No-Asymmetries Argument establishes that, within the practice of reasoned discourse, rationality compels players to assert their from their stock of warranted beliefs. A person cannot genuinely take on the role of a player in a game of reasoned discourse and not be under defeasible rational pressure to sincerely assert their warranted

beliefs. This rational pressure might, and often does, get thwarted by the particularities of a person's given situation. Nevertheless, this rational pressure is never fully eliminated. Instead, it constitutes the "first word" in what a person may do in any particular game of reasoned discourse. What strategy a person pursues is informed by whether they're in a scenario where they're entitled to presume no relevant epistemic or cognitive advantage over their interlocutors. That entitlement is modulated as the game evolves.

In addition to Rescorla's No-Asymmetries argument, the other consideration in support of resisting doubt to GD5 is the observation that we are not simply rational agents. We are socially rational agents. While there might be some kinds of rational agents for whom practices of conversing and reasoned discourse are purely accidental or contingent developments, we are not those kinds of agents. As socially rational agents, specifically linguistically socially rational agents, conversational practices, including the practices of reasoned discourse, are essential facets of our form of agency.

You might ask what it would be for a person to be an essentially socially rational agent. This is admittedly difficult for me to articulate. Here is my tentative initial gloss on what it is to be an essentially social rational agent. First, I take for granted that socially rational agency, like other forms of agency, is essentially a capacity for goal oriented behavior. This is what our form of agency has in common with other forms of agency. For example, consider the agency of animals whose sensory capacities fall short of representational.<sup>49</sup> By this, I mean that at minimum, explanatory generalizations regarding their sensory capacities and how those capacities relate to behavior need not invoke any semantic kinds. Semantic kinds include contents, such as modes of presentation, and satisfaction conditions. Brute causal correlations between states of sensory stimulation and behavior are all that's needed. Something like this type of agency is displayed by such animals as hydra and earthworms. Hydra and earthworms don't represent, in the sense

---

<sup>49</sup>This is the sort of agency that Burge calls "primitive agency". See his (2009b) and (2010), ch 8.

that genuine explanations of their behavior invoke contents or satisfaction conditions.

Second, I take for granted that socially rational agency is a species of representational agency. By this I mean that socially rational agency is a kind of agency that involves behavior that's oriented around the (semantic) representation of some goal. This is the sort of agency that we share with creatures, such as honeybees and other arthropods. In addition to merely sensory capacities, arthropods have representational sensory capacities, i.e., perceptual capacities. Such animals orient their behavior along the perceptual representations of goals. Since these capacities are representational, they are individuated, at least in some part, by principled relationships that go beyond states purely internal to the individual. They are individuated, at least in some part, by principled relationships between their content bearing states and the purported subject matters of those states. For example, a honeybee's behavior of collecting honey from flowers is explained, at least in part, by their capacity for perceiving ultra-violet markings.

Third, I take for granted that rational agency, be it social or merely individual, is a species of conceptual agency. By this, I mean that our behavior is mediated, at least in part, by not only perceptual representations of a goal, but by reasoned plans or intentions. These reasoned plans are essentially complexes of propositional attitudes. Paradigmatic examples of propositional attitudes include beliefs, desires, intentions, hopes, wishes, presuppositions, and so on. These propositional attitudes are mental states that consist in an individual taking a certain "stance" or "attitude" towards a proposition. These propositions are, at minimum, logically and quantificationally structured semantic contents. They are logically and quantificationally structured in the sense that they are formed not only via the application of first order predicate functions to terms, but also by the recursive application of logical operators to fully formed propositions and by the application quantifiers to first order predicate functions.

Fourth, as I explained in an earlier chapter, well functioning rational agency involves



unifying the aspects of reason. This is clearest when considering the difference between a person who promotes the practical aspects of reason over the impersonal aspects versus a person who promotes the practical and impersonal aspects together. As exemplified by many different accounts of impersonal rationality, that advantage consists in a systematic protection or robustness from certain types of risks. For example, a person who unifies the aspects of reason enjoys safety from error<sup>50</sup>, enjoys immunity from guaranteed loss<sup>51</sup>, or is less susceptible to decisive challenges or counterexample in an idealized game of reasoned discourse.<sup>52</sup>

Fifth, and admittedly controversially, an individual whose type of rationality is social rationality whose well-functioning requires unification of not only the aspects of reason as applied to their individual propositional attitudes. Social rationality requires that whenever possible, a person unifies the aspects of reason as applied to their individual attitudes together with the aspects of reason as applied to their social behaviors. I think that there are intuitively two dimensions to this requirement. The first dimension is straightforward: to the extent that a socially rational agent is able to exercise their social capabilities, that's the extent to which they must also unify all the applicable aspects of reason. This covers, for instance, people whose social capabilities are neither divergent nor impaired but who have been situationally robbed of new opportunities for exercises of social reasoning. This would be the case of a person who is stranded on a remote island with slim to no prospects for seeing another person.

The second dimension is what I think is a prototype or ancestor of moral agency, if not plainly moral agency itself. Social agency requires, among other things, that members of a social group do what they can to promote each other's social agency. This means, among other things, that there's some minimal defeasible rational pressure for social agents to

---

<sup>50</sup>Williamson (Williamson2000)

<sup>51</sup>Ramsey (1926/2011)

<sup>52</sup>Rescorla (2007)

partake in shared activities, such as conversation and reasoned discourse. Among other things, this also means that there's some minimal defeasible rational pressure for social agents to accommodate each other's limitations when partaking in shared activities.

Drawing inspiration by Kant (1785/2018), I present this as a standing general commitment for people to "respect the humanity of others"—i.e., to acknowledge and promote the social rational agency of other people. Among other things, this suggests that persons whose realizations of the capacities for social reason are either divergent or impaired should do their best to unify all the aspects of reason to the extent that they're able. But more importantly, it means that socially rational agents must do their best to accommodate and promote unified individual and social reasoning in others. This means that socially rational agents must strive, to the extent that they're able, to accommodate and promote unified individual and social reasoning in people whose social capacities are divergent or impaired.

Given the five considerations offered above, I maintain that socially rational agency differs from merely individually rational agency in the following way. Socially rational agency generally involves the ability to socially reason, i.e., engage in shared activities, including the practices of conversation and reasoned discourse.<sup>53</sup> Socially rational agency also means that a person's individual and social reasoning is subject to a meta-aspect: the aspect of promoting unified social and individual reasoning in others. I call this meta-aspect the basic social aspect of reasoning.

---

<sup>53</sup>Strictly speaking, we might perhaps call this discursively rational agency. For we might imagine forms of socially rational agency that involve a generally social element, but of a form that's more rudimentary and doesn't involve language in the sense required for conversation and reasoned discourse. I suspect that many kinds of animals have variations on this alternative form of socially rational agency. For example, I suspect that animals with social patterns of behavior might have variations or prototypes of this non-discursive socially rational agency. The animals that come to mind as potential examples involve animals that seem to have display behaviors like tool use and culture. Such animals include certain types of corvids, non-human primates, and elephants. I concede that beyond the mere conception of this non-discursive form of socially rational agency, it's an empirical question whether there are any animals that actually have this kind of agency. For the sake of simplicity, I set aside this important wrinkle.

You might think that my appeal to a general social aspect of reasoning commits some problematic hyper-intellectualization of socially rational agency. An account of the mental problematically hyper-intellectualizes a psychological phenomenon when the conditions for attributing that psychological phenomenon to an individual are too psychologically demanding. Applied to my discussion, the complaint is that my account of socially rational agency requires that a person be able to recognize socially rational agency in other people. This complaint seems plausible insofar as we might think that a person's promoting the basic social aspect of reasoning involves them being able to conceptualize such complex, philosophically sophisticated concepts as socially rational agency, individual reason, social reason, and being unified in reasoning.

I'm not deterred by this hyper-intellectualization worry. Among other things, it doesn't seem to me that being subject to an aspect of reasoning necessarily requires that the person be able to recognize any philosophically sophisticated concepts as socially rational agency, individual reason, social reason, and so on. Rather, all that matters is that the person enjoys some base level sensitivity such phenomena. This sensitivity could even be indirect. But this sensitivity does not require actually recognizing anything. A person who possesses genuinely socially rational agency in the relevant sense need not be in a position to think that they must promote individual and social reasoning in others. All that matters is that in some situations, if the person merely senses or detects some considerations that indicate or are causally correlated with social agency in some form or other, that this person's is disposed to take actions that would contribute to the promoting that social agency's being unified.

I understand that the view that we are merely individual rational agents holds a lot of cultural sway among philosophers. I think that these ideas find sway, even among ostensibly social accounts of reasoning. For example, there is a social contract tradition

in moral and political philosophy.<sup>54</sup> This tradition has even recently crossed over to epistemology.<sup>55</sup> This social contract tradition often presents social endeavors as things that merely individually rational agents, in ideal conditions, can either ratify or reject. On such a presentation, socially rational agency might seem to be something that is not genuinely its own form of agency, but rather something that contingently emerges when enough merely individually rational agents, in ideal conditions, engage in certain behavioral patterns that emulate the determining of a convention.

Another example of this individualistic view of our rational agency is presented in Hegel's *Phenomenology of Spirit*.<sup>56</sup> There, Hegel provides what could be described as a genealogy of rational agency. There, agents are described as developing from merely sensory beings, to perceiving beings, to conscious beings, and then to self-conscious beings. Self-conscious beings only become so when they finally encounter other conscious beings and are forced to submit into "bondage" in order to preserve their own existence. Again, this suggests an individualistic perspective on rational agency. The idea here is that socially rational agency is a level of development that a person can achieve. However, such achievements might be read as contingent developments of an individual agents rather than as a way in which such an agent is functioning well.

Yet another example of this individualistic view is present even in the Kantian moral philosophy that serves as my inspiration. In particular, the way that Kant presents

---

<sup>54</sup>The contemporary entry point into this tradition is Rawls (1955)

<sup>55</sup>For the contemporary entry point into this development, see Craig (1990). Simion (2021) offers a discussion of this perspective as it applies to the epistemology of testimony. Simion's discussion is, in many ways, broadly congenial to the discussion I'm offering now. It purports to offer an *a priori* defense of ambitious presumptivism. It also purports to appeal to the view that there is something essentially social about knowledge and warrant by way of a social contract theory of knowledge and a knowledge-first account of warrant. Nevertheless, the motivations for Simion's discussion is the presumption that the challenges to Burge's functional unity argument are decisive. That presumption is precisely what I resist in this essay. Nevertheless, it's possible for one to use a social contract analysis of knowledge or warrant to establish the points that I'm attempting to establish now. I leave this as an exercise for another occasion.

<sup>56</sup>Hegel (1977)

his tests for the famed categorical imperative suggests a fundamentally individualistic perspective on rational agency.<sup>57</sup> The tests for the categorical imperative often seem to take the form of thought experiments performed by an individual. The person entertains whether a candidate course of action is one that can be willed under certain restrictions. The morally worthy will would only endorse as a course of action those candidates that can be will under those restrictions. The morally unworthy will would pass actions that fail to satisfy those restrictions. This methodology can give off the impression that social rationality ultimately bottoms out in individual rationality. After all, Kant's tests for the categorical imperative seem to frame what's socially rational in terms of what constraints there are on an individual's reasoning.

Nevertheless, I think that it is a mistake to think that the view that we are essentially socially rational agents is unprecedented. In the cases of social contract theory, Hegel's account of consciousness and self-consciousness, and Kant's tests for the categorical imperatives, the suggestions of a fundamentally individualistic perspective strike me as superficial. In its core applications the social contract tradition is meant to give an ideal model of what it is for a state to be politically legitimate. It's not meant to actually speak to what kinds of rational agents we are. To the extent that it says anything about the nature of socially rational agency, it suggests that such agency has a structure akin to a legislature of rational agents.<sup>58</sup>

Similarly, the genealogy of rational beings given in Hegel's *Phenomenology* is meant to distinguish different kinds of agency in a way that is narratively compelling. It need not be taken as a statement that we all start off as merely sensory agents who must proceed through stages of developing perception, consciousness, and then finally self-

---

<sup>57</sup>Kant (1785/2018)

<sup>58</sup>This imagery abounds in some applications of social contract theory to morality. Fo a particularly prominent example of this trend, see Korsgaard (2009) who attempts to analogize the rationality of an individual person in terms of the proper arrangement of a politically legitimate state.

consciousness. Setting aside the empirical issues with taking Hegel's discussion literally, the overarching lesson from Hegel's discussion is that social rationality involves a sensitivity to the agency of others. This sensitivity manifests in how, when people engage in shared activities and other social endeavors, people have to subordinate themselves to some sort of lord or sovereign. There's no reason to think that this lord or sovereign is not just an expression of their shared reasoning.<sup>59</sup>

Also in this vein, Kant's moral philosophy also ultimately belies a view that our form of rational agency is ultimately social. After all, for Kant, the categorical imperative is a general standard on our form of rational agency. Further, the most comprehensive and attractive formulation for the categorical imperative, kingdom of ends formulation, introduce social elements. The kingdom of ends formulation suggests that the test for the moral worthiness of a candidate course of action is whether it could be willed as a law by a legislature of ideally rational beings in a community where everyone respects and promotes the humanity in each other.<sup>60</sup> This suggests that Kant simply takes for granted that there is an irreducibly social aspect to our form of rational agency. It is irreducibly social because moral concerns are social concerns. Finally, these moral concerns are at the heart of Kant's account of our form of rational agency. Thus, it seems that Kant assumes that we are socially rational agents and leverages that assumption to help articulate an account of what it would be for a person's action to be morally worthy.

Setting aside historical entries in moral philosophy and social contract theory, individualistic perspectives in contemporary philosophical theorizing on agency and the mind

---

<sup>59</sup>Of course, this hastily sketched discussion avoids the details of most deep Hegel scholarship. However, my discussion's focus on Hegel's account of self-consciousness, and the importance of the lord-subordinate relationship, is I think supported by much of the mainline Hegel scholarship. That focus began with Kojève (1969). If contemporary Hegel scholarship is any indication, Hegel's account of social reasoning might be, for the present purposes, hyper-intellectualized after all. That's because Hegel does seem to think that social reasoning's intrinsic purpose is to comprehend itself and its place in the world. This sounds like For an excellent overview of Hegel scholarship, see Redding (2020)

<sup>60</sup>Kant (1785/2018) 4:439

are fairly plentiful. I think these views are best described as instances of what Burge calls “individualist representational” views of mental states.<sup>61</sup> Examples of individualist representationalism include sense data theories of perception, such as those advanced by H.H. Price and Rudolph Carnap.<sup>62</sup> They also include the descriptivist theories of mental and linguistic reference, such as those advocated by P.F. Strawson and W.V.O. Quine.<sup>63</sup> On these views, mental states are not constituted by principled connections between their contents and the purported subject matters of those contents. Instead, mental states are constituted by relationships that must be established by the individual.

I think it’s plausible to say that individualist accounts of mental and linguistic representation are no longer the dominant accounts. This is due in part to the advent of modern anti-individualist accounts of mental and linguistic representation. In particular, many of the influential defenses of anti-individualism about mental and linguistic representation bring in irreducibly social considerations. For instance, Kripke’s positive account of the meaning of proper names introduces the idea that the meanings of proper names are not exclusively fixed by a person’s beliefs about the bearer. Rather they’re fixed in part by a social history of use that links the person’s use of that name to an initial “baptism” wherein the name was first applied to the purported referent.<sup>64</sup> Kripke applies similar lessons to natural kind terms.<sup>65</sup> Putnam defends a similar thesis regarding natural kind terms by reflecting on how scientific social practices establish a division of labor. This division of labor fixes the purported references of natural kind terms and then disseminates those references by way of how scientific practice parcels out aspects of scientific inquiry across the broader community.<sup>66</sup> Finally, Burge’s most

---

<sup>61</sup>Burge (2010).

<sup>62</sup>Price (1964); Carnap (1967).

<sup>63</sup>Strawson (1959); Quine (1960)

<sup>64</sup>Kripke (1980)

<sup>65</sup>ibid.

<sup>66</sup>Putnam (1975)

influential contribution to this tradition establishes analogous claims regarding the contents of propositional attitudes by reflection on how, in counterfactual scenarios, different patterns of use established by the scientific community lead to a person having genuinely different thoughts and behavior.<sup>67</sup>

Anti-individualism about mental and linguistic reference illuminate how a person's mental states and linguistic competencies are determined in part by a person's broader environment, including their broader social environment. Insofar as anti-individualism is a theory about what it is for us to have mental states, the social considerations marshaled by anti-individualism's proponents also lend credence to the idea that our form of agency is irreducibly social. They do so because they show that many of the capacities we deploy as rational agents are determined by factors in our broader environment. These factors include social factors.

Anti-individualism about mental and linguistic reference isn't the only contemporary precedent for a thesis about mind and language on which we are social creatures. For example, Brandom's pragmatist account of mental and linguistic contents is contributes to this precedent.<sup>68</sup> For Brandom, reflection on mental and linguistic contents takes, as its starting point, the observation that we are essentially socially rational agents. As he sees it, the evidence base for a theory of linguistic meaning and the contents of our propositional attitudes must take, as its starting point, our communicative practices. Among these communicative practices, our practices of reasoned discourse are explanatorily privileged. The result is a proposal on which our individual reasons, e.g., our beliefs, desires, and intentions, are essentially social because those mental states get their structure and content from our practices for reasoned persuasion and reasoned inquiry.

Arguably, Brandom's pragmatist account of mental and linguistic contents is really

---

<sup>67</sup>Burge (1979)

<sup>68</sup>Brandom (1994)



a particularly extreme version of anti-individualism. In particular, it seems to be an instance of what Burge (2010) calls “deflationary anti-individualism”. Deflationary anti-individualism is a species of anti-individualism that affords no substantive explanatory role to semantic kinds such as contents or satisfaction conditions. Brandom’s own elaborations of his own arguments confirm this agenda.<sup>69</sup> At the same time, Brandom’s proposal is also problematically hyper-intellectualized. This is because, for Brandom, all mental states, including sensory perceptions, get their rational structure and content from their place in our practices of reasoned discourse. He holds the same for all agency. On his view, all action gets its rational structure, and whatever contents they might bear, from their place in our practices of reasoned discourse.

I think that the faults with Brandom’s account can be mitigated by adjusting our expectations of what it can reasonably establish. A full development of this proposal would, all by itself, be a weighty endeavor. However, the main idea is fairly simple. Let’s not aspire to give an account of agency, language, and mind on which satisfaction conditions do not play a central role. We can then let some other version of anti-individualism explain the fixation of content. Further, let’s not aspire to assimilate all agency under the heading of socially rational agency. Instead, we should simply admit for ourselves that all animals, ourselves included, exhibit multiple forms of agency. Socially rational agency is but one of the forms of agency that we exhibit. From there, we set for ourselves the much more modest goal of explaining how socially rational agency is structured. This means that the aspects of our agency that are genuinely social are structured in the way that Brandom has in mind. Specifically, our rational agency, to the extent that it is social, is structured in the ways that Brandom has in mind.<sup>70</sup> And like Brandom presumes,

---

<sup>69</sup>Brandom (2000)

<sup>70</sup>With caveats, of course! As Pryor (2005) and Rescorla (2009b) forcefully argue, the structure of warrant is not necessarily as dynamic as a game of reasoned discourse. I think this is plausible because, like the practical aspect of reason, the impersonal aspects of reason are importantly prior to the social aspect of reason. Insofar as warrant is within the purview of the impersonal aspect of reason, we should

our form of agency essentially involves socially rational agency. Unlike Brandom, we admit that our form of agency also involves other aspects of agency, including perceptual agency.

I submit that many currents in moral and political philosophy, as well as in the philosophy of language and philosophy of mind, provide precedence for the idea that we're socially rational creatures. This means that even though I don't have any novel account for why we are socially rational agents, the presumption that we are socially rational agents is not particularly philosophically novel or outrageous. Indeed, if the presumption's ubiquity in moral and political philosophy, as well as in recent anti-individualist accounts of mental and semantic content are to be taken seriously, the presumption that we are socially rational agents is a common sense presumption. If it is plausible that we are socially rational agents, then it is plausible that we must unify all our aspects of reason, in both their individual and social deployments.

That concludes my attempt to address doubt for GD5. GD5 was the crucial claim that if belief and assertion both fall under the jurisdiction of an impersonal aspect of reason, then reason functions in one mind so as to present truth to another mind. you might doubt GD5 because you might suspect that there's a gap between our rational agency and the practices of reasoned discourse. In response, I marshaled two considerations against the force of that doubt. The first consideration was a version of Rescorla's No-Asymmetries Argument. The conclusion of that use of the argument was that our practices of reasoned discourse generated default rational pressure to be unified, i.e., to assert what you genuinely believe to be true. The other consideration was the intuition that we are essentially socially rational agents. If we are socially rational agents, then we must unify the impersonal and practical aspects of reason not only within individual

---

expect that the structure of warrant has somewhat different properties from the structure of reasoned discourse.

and social reasoning reason, but across individual and social reasoning. That is, we must unify the impersonal and practical aspects of reasoning not only within our own minds and within the practices of reasoned discourse, but across reasoned discourse. If both of these considerations are plausible, then it's also plausible that reason functions in one mind so as to present truth to another. Hence, a rational agent, at least of the sort that we are, both believes and asserts the truth.

## 5.5 Some Closing Remarks

In addition to rescuing Burge's functional unity argument, the two considerations in support of our social agency also suggest a straightforward explanation of how it is that a person's assertions defeasibly indicate their beliefs. A person's assertions defeasibly indicate their beliefs because in an idealized game of reasoned discourse a person's default rational strategy is to assert what they believe.

This consideration predicts that the correlation between what a person asserts and believes is defeasible. On this consideration, the correlation between what a person asserts and what person believes is defeasible because in some particular instances of a game of reasoned discourse, a person might be called to assert something they don't believe. For example, in a game of reasoned discourse, the players might be reasoning their way through an instance of a *reductio ad absurdum*. In a social realization of a *reductio ad absurdum* you justify the assertion of a conclusion by assuming and reasoning from its negation. This essentially involves asserting something you don't believe.

Here is another way in which, within a game of reasoned discourse, the correlation between what a person asserts and what a person believes could break down. In a particular game of reasoned discourse a speaker might find it situationally advantageous to lie. In lying, they essentially asserting something they don't believe. They might

understand that it is riskier to lie. Nevertheless, they might have reason to believe that they possess some epistemic or cognitive advantage over other participants in the discourse. Such a person might not be unified in their reasoning. But unification is a standard, not a causal law. Reasoners fall short of the standard of unification all the time.

Another reason a person's assertions defeasibly indicate their beliefs is that a person, in the relevant sense, is not just a rational agent but a socially rational agent. As a socially rational agent, they must promote unified functioning within and across individual and social reasoning. Accordingly, a person must believe what it is they assert. They must also, when discursively appropriate, assert what they believe.

Even on this consideration, the correlation between what a person asserts and what they believe can break down. This happens because unification is hard. A person might be tempted by the prospects for advancing their plans by presenting the truth to their own mind but distorting the truth in others. That is to say, it is often times the case that a person can advance their own personal projects by asserting lies to others. In such cases, there is understandably a temptation to assert lies. And at least some of the time, a person acts on that temptation. Such a person is disunified as a socially rational agent. But such disunity is not unimaginable. It is a consequence of the very ordinary drive to achieve one's own goals.

I think it's important to maintain these brief reflections on how my development of a Burge-Glanzberg account can explain the defeasible correlation between what a person asserts and what they believe. I think it's important because the rival belief-expression view does so. The rival belief expression view would undermine Burge's functional unity argument if true. But an attraction of the belief-expression view is that it respects the idea that a person's assertions defeasibly indicate their beliefs. If the Burge-Glanzberg account is to be plausible, it needs to be able to explain that defeasible indication. I con-

tend that these brief closing reflections, combined with my final response to the rejoinder against my reply to the fifth objection to the functional unity argument, establish the on-balance plausibility of the functional unity argument.

## Chapter 6

# Burge, Kant, and the Murderers at the Door

Over three rounds of objections and replies, I have defended the plausibility of both the functional unity argument and my interpretation of Burge as committed to that argument. In doing so, I explicated a Veridicality account of Rationality. I have also drawn complimentary relationships between the functional unity argument, the truth goal account of assertion, the reasoned discourse account of assertion, and the plausible presumption that we are socially rational agents. It should be clear that I think that Burge's functional unity argument is an underappreciated development in the defenses of ambitious presumptivism. It should also be clear that I think that Burge's functional unity argument, once supplemented with an appropriately sophisticated account of assertion, offers a promising template for examining *discursively socially* rational agency, i.e., the socially rational agency of creatures like us who are able to engage in conversation and reasoned discourse.

In this closing chapter, I consider how Burge's functional unity argument and the resulting Veridicality Account of Rationality can be defended against one final objection. This final objection brings us back to my initial comparison of Burge's functional unity argument with Kant's explanations of the categorical imperative. When I initially in-

roduced Burge's functional unity argument, I cited as a historical predecessor Kant's explanations of the categorical imperative. I thought this comparison was apt because both Burge and Kant seem to be arguing that a person who asserts lies rather than truths is irrational, at least in the ideal sense of rationality. We might plausibly say that the functional unity argument constitutes Burge's contribution (intentional or not) to a Kantian account of the rationality of lying. The Kantian account of the rationality of lying is that lying is irrational.

The type of objection that I want to consider is what I call a murderer-at-the-door objection. Because of the troubles such objections caused Kant, murderer-at-the-door objections strike me as a good benchmark for Kantian accounts of the rationality of lying. If a Kantian account of lying is to constitute an improvement over the original, it needs to offer a satisfying explanation of the rationality of lying in such cases.

## 6.1 The Murderers At The Door

Here's a classic murderer at the door style case:

### **LYING TABITHA**

Mary is being pursued by Josephine, who plans to murder her. Mary escapes to Tabitha's house and seeks refuge, which Tabitha grants. Josephine, who has lost track of Mary, goes door to door and eventually makes it to Tabitha's house. Tabitha answers the door and Josephine asks where Mary has gone. Tabitha knows about Josephine, but Josephine doesn't know that Tabitha knows. Tabitha also knows that if she keeps silent, Josephine would be able to infer that something is up and perhaps turn her hostilities towards her. Tabitha lies, saying that she saw Mary heading further down the road towards the center of town. Josephine accepts this and heads off towards the center

of town. Mary thanks Tabitha and steals away towards the outskirts of town.

Tabitha's lie seems to be both *prima facie* on-balance moral and *prima facie* on-balance rational. Tabitha's lie seems so moral because protecting people in danger seems to be the morally correct thing to do. Tabitha's lie seems so rational because doing so would serve her moral end protecting people in danger.

Here's another murderer at the door case, based on innovations due to Seanna Shiffrin:<sup>1</sup>

### LYING AGATHA

Mary is being pursued by Josephine, who plans to murder her. Mary escapes to Tabitha's house and seeks refuge, which Tabitha grants. Josephine, who has lost track of Mary, goes door to door and eventually makes it to Tabitha's house. Tabitha answers the door and Josephine asks where Mary has gone. Agatha, Tabitha's next door neighbor, has noticed what's going on, and she knows that none of the involved parties have noticed her snooping. She knows that she can feign ignorance, and even absence without repercussion. Nevertheless, she decides to pop her head out her window and say something before Tabitha can say anything. She lies, saying that she saw Mary heading further down the road towards the center of town. Josephine accepts this and heads off towards the center of town. Mary quietly thanks Tabitha and Agatha and steals away towards the outskirts of town.

Shiffrin thinks that we should have the same verdicts about Agatha in the LYING AGATHA case as we do about Tabitha in the LYING TABITHA case. Again, it seems morally proper to protect people in danger. Agatha noticed that Tabitha and Mary were

---

<sup>1</sup>(2014), p. 30.



in danger to and she took steps to protect them. Given her moral ends, it seems rational for Agatha to lie if she intervenes.

In both of these cases, it seems that the rational thing for a person to do is to lie. In both of these cases, it seems as though lying is the rational thing to do because the lies would promote the liar's intended ends. Further, these intended ends seem to be morally correct ends.

## 6.2 Kant and the Murderers at the Door

Historically, Kant's account of the irrationality of lying is difficult to reconcile with our intuitions about Murderer-at-the-door cases. In this section, I briefly recap Kant's discussion of the irrationality of lying. I then explain how this conflicts with our intuitions about the murderer-at-the-door cases above.

Kant most directly addresses the case of lying when, in the course of elaborating on the Universal Law Formulation of the Categorical Imperative, he discusses the case of lying promises. Let's start by considering Kant's application of the Universal Law Formulation of the Categorical Imperative. On that formulation, the Categorical Imperative tells against willing any maxim (i.e., any potential plan of action) that could not be willed as a universal law of nature. Intuitively, a maxim that could not be willed as a universal law of nature is one that, intuitively, would lead to some kind of break down or "inconsistency".

In applying the Universal Law Formulation of the Categorical Imperative to the question of whether lying is rational, Kant writes:

...would I indeed be content that my maxim (to get myself out of difficulties by a false promise) should hold as a universal law (for myself as well as for others)? and could I indeed say to myself that everyone may make false promise when he finds himself in a difficulty he can get out of in no other

way. Then I soon become aware that I could indeed will the lie, but by no means a universal law to lie; for in accordance with such a law there would properly be no promises at all, since it would be futile to avow my will with regard to my future actions to others who would not believe this avowal or, if they rashly did so, would pay me back in like coin; and thus my maxim, as soon as it were made a universal law, would have to destroy itself.

Kant (1785/2018) 4:403

Here, Kant acknowledges that there are two ways to take the question of whether lying is rational.<sup>2</sup> The first way to take the question of whether lying is rational is to take it as a question of what a person can prudently will. The maxim of lying to get out of a difficult spot is certainly something a person can will. Further, the course of action described in the maxim is something that, depending on circumstances, could be prudential in the sense of being self serving.

The second way to take the question of whether lying is rational is to take it as a question of ideal, on-balance rationality. In this sense, Kant denies that lying is rational. Kant denies that lying is rational in the sense of ideal on-balance rationality because of the Universal Law Formulation of the Categorical Imperative. It isn't rational in this deeper sense because, if everyone were to adopt the maxim as a matter of universal law, the practices of assertion and promising would break down. They would break down because the connection between the world and what is asserted or promised would break down. Without the connection between the world and what is asserted or promised, the practices of assertion and promising would no longer be valuable.

In some ways, Kant's application of the Universal Law Formulation of the Categorical Imperative is a loose precursor to the 20th century Kant-inspired transcendental arguments against the unreliability of assertion. Two major advocates of such transcendental

---

<sup>2</sup>Kant (1785/2018) 4:401-402

arguments against the unreliability assertion are Coady (1992) and Davidson (2001b). Here's the brief unifying gloss on these arguments. The first premise is that we actually a practice of asserting. The second premise is that a necessary condition of the possibility of such a practice is that assertions are in fact, at least in the long run, reliable. The third, premise is that the actuality of the practice of asserting entails its possibility. Thus, assertions are in fact, at least in the long run, reliable. Arguably, Kant's discussion is a precursor to this kind of argument because he thinks that the unreliability of assertion that results in lying being willed as a universal law is incompatible with the possibility of the practice.

Kant's discussion of the Humanity Formulation of the Categorical Imperative also points towards the same conclusion that the Categorical Imperative tells against the rationality of lying. The Humanity Formulation of the Categorical Imperative goes like this. On this formulation, we must never in such a way that we treat humanity, in ourselves and in others, as a mere means rather than an end to be promoted. Accordingly, if a candidate maxim simply exploits a person's humanity without "respecting" it, then it is not the sort of maxim that a rational person could adopt. In this case, respecting a person's humanity involves acknowledging their status as a socially rational agent and not undermining it.

When Kant applies the Humanity Formulation of the Categorical Imperative to the question of lying, he writes:

...he who has it in mind to make a false promise to others sees at once that he wants to make use of another human being merely as a means, without the other at the same time containing in himself the end. For, he whom I want to use for my purposes by such a promise cannot possibly agree to my way of behaving toward him, and so himself contain the end of this action.

This conflict with the principle of other human beings is seen more distinctly if examples of assaults on the freedom and property of others are brought forward. For then it is obvious that he who transgresses the rights of human beings intends to make use of the person of others merely as means, without taking into consideration that, as rational beings, they are always to be valued at the same time as ends, that is, only as beings who must also be able to contain in themselves the end of the very same action.

Kant (1785/2018) 4:430-431

In parsing this passage, I follow Korsgaard's intuitive explanations.<sup>3</sup> According to Kant, asserting a lie is not something that can pass the Humanity Formulation of the Categorical Imperative. Asserting a lie action doesn't pass the Humanity Formulation because asserting a lie essentially involves disrespecting humanity in others. The person being lied to cannot genuinely participate in the exchange we're having together because he is being lied to. The liar has obfuscated the real purpose of the interaction. The person being lied to is essentially being coerced. The fundamental wrong, then, of lying is that in lying, the liar has forced their victim into an activity for which they did not freely consent. It is for this reason that lying does not pass the Humanity Formulation. It is also the reason for why lying is always morally wrong, regardless of a person's intentions.

Kant's two applications of the Categorical Imperative to the case of lying point towards the thesis that lying is always morally wrong. Given the status of the Categorical Imperative in Kant's account of rationality, we can say that lying is always morally wrong because lying does not pass a crucial general standard for rational action. That standard is the categorical imperative. Lies are always morally wrong because they are not the sorts of maxims that a morally worthy reasoner could adopt. This is so, according to Kant, because if every one were to lie, practice of asserting would lose all value. The

---

<sup>3</sup>Korsgaard (1996) pp. 36-37

practice of asserting would lose all value because no one would want to engage in it. No one would want to engage in it because they would not want to be coerced into activities they did not freely consent to.

Kant's treatment of lying seems to conflict with our intuitions about the **Lying Tabitha** and **Lying Agatha** cases. In those cases, our intuition is that the liars in those cases are both rational and moral in doing so. In those cases, if the liar were to tell the truth, they would become contributors to what is likely to be an immoral action. This leads to a tension for Kant's treatment of the rationality of lying. On the one hand, lying violates the Categorical Imperative. Since the Categorical Imperative, for Kant, is one of the most general rational norms, it follows that lying is generally irrational (and thus, immoral). On the other hand, it seems as though a person who would tell the relevant truth instead of the lie in murderer-at-the-door cases would contribute to an immoral action. And immoral actions, for Kant, are actions that do not pass the tests for the Categorical Imperative. Since these actions are violations of the Categorical Imperative, it follows that, for Kant, such actions are also irrational. So, lying in a murderer-at-the-door case is irrational. But so to is telling the truth.

You might think that if both lying in a murder-at-the-door case and telling the truth in a murderer-the-door case is irrational, then doing neither would be the the rational option. This idea might seem attractive because it would seem to constitute a workable middle ground. It might seem like an attractive middle ground because it allows us to agree with both sides of the issue while at the same time offering a positive proposal about how to proceed.

Nevertheless, I think that in the context of a murderer-at-the-door case, the action of neither asserting the truth nor of asserting a lie is also going to be irrational in Kant's sense. For consider what would happen if everyone adopted, as a matter of universal law, the maxim of neither asserting the truth nor asserting a lie. In such a scenario, it again

seems as though the practice of asserting would break down. The practice of asserting would break down because the practice requires people to assert propositions pertinent to the issue under discussion. If people asserted neither sincerely nor insincerely, they would either not assert at all or else assert on irrelevant propositions. But this would just be to stonewall. Either way, the practice of asserting would break down fairly quickly. For if participants either made no assertions or else stonewalled with their assertions, there would be no value in participating in the practice of reasoned discourse. There would be no value in the practice of reasoned discourse because under such a universal law, the intrinsic goals of reasoned discourse are never achieved.

Depending on which variant of the murderer-at-the-door cases one considers, the action of asserting neither truths nor lies also seems to fail the humanity formula. An example of such a cases is the **Lying Tabitha**. In that case, the actions of asserting neither truths nor lies invites the possibility of harm upon oneself. Here is why. Failing to answer the questions of a murderer in pursuit of an immoral end is liable to make that murderer angry. In making the murderer angry, you have no guarantee that they won't turn their murderous designs on to you. If you draw their ire in this way, you're not respecting the humanity in yourself. This is because respecting the humanity in yourself calls for you to strive towards self-preservation. Stonewalling a murderer in pursuit of their next kill does not contribute towards self-preservation.

Hence, I submit that Kant's strong stance on the rationality of lying generates trouble for him. It generates trouble for him because it seemingly engenders the disturbing result that there's no rational thing for a person to do in a murderer-at-the-door case. But this is counter to our intuitions about murderer-at-the-door cases. In such cases, there seems to be a rational thing to do. In such cases, the rational thing to do seems to be asserting a lie.

### 6.3 Burge and the Murderers at the Door

Just as Murderer-at-the-Door cases pose a problem for Kant's ethics, they also pose a problem for Burge's account of agency. In this section, I explain why it's difficult. In subsequent sections, I sharpen the difficulties further before finally offering a promising reply on Burge's behalf.

Let's consider what Burge would say about the **Lying Tabitha** case. In that case, our intuitive verdict is that it is rational for Tabitha to assert a lie in that case. Let's consider whether Burge's view can recover this verdict. Is it in fact, on-balance generically rational for Tabitha to lie in that case? At best, Burge can say that it is *prima facie* on-balance instrumentally rational for Tabitha to assert a lie. He can say so because, for Tabitha, lying would satisfy her interests in bringing about a moral outcome.

Once we ask whether Tabitha is generically rational, it seems that Burge has to say no. It appears he has to say no because lying occasions a disunity in the function of reason. Accordingly, in the case of lying, the instrumental aspect of reason is promoted at the expense of the truly transpersonal aspect. This disunity undermines the rationality of Tabitha's lying, in that occasion. It does so because it defeats the on-balance generic rationality of any lying on any occasion.

What Burge would say about the **Lying Tabitha** case, he would also say about the **Lying Agatha** case. The best that Burge can do is grant the *prima facie* on-balance instrumental rationality of asserting a lie. Agatha of the **Lying Agatha** case could satisfy her goals of supporting the moral outcome by asserting lies rather than truths.

Again, once we switch from the question of instrumental rationality to the question of generic rationality, Burge has to say that it's not on-balance generically rational for Agatha or Tamara to assert lies rather than truths. He has to say no in these cases because lying occasions disunity in the function of reason; the instrumental aspects of

reason are promoted at the expense of the truly transpersonal aspect. This disunity undermines the generic rationality of either Agatha or Tamara lying in their respective circumstances. It does so because it undermines the on-balance generic rationality of any lying on any occasion.

These considerations suggest the beginnings of an objection to Burge's functional unity argument. In its simplest form, the argument looks like this:

**MDO1** If the functional unity argument is sound, then it is not true that in the **Lying Tabitha** and **Lying Agatha** cases, it is on-balance generically rational for a person to assert lies rather than truths

**MDO2** Nevertheless, it is true that in the **Lying Tabitha** and **Lying Agatha** cases, it is on-balance generically rational for a person to assert lies rather than truths

**MDO3** So, the functional unity argument is unsound

I call this argument the Murderers-at-the-Doors Objection. MDO1 is justified as a logical consequence of the functional unity argument. MDO2 is supported by ordinary intuitions about the **Lying Tabitha**, **Lying Agatha**, and **Lying Tamara** cases . MDO3 follows from MDO1 and MDO2 by and *Modus Tollens*.

The Murderer-at-the-Door Objection is not, as pitched, the most specific objection to the functional unity argument. It does not, as formulated, tell us which of the six premises of the functional unity argument fails. Nevertheless, it provides us the mental resources to nevertheless doubt the soundness of the argument. It does so by independently supporting the negation of the conclusion of the functional unity argument.



## 6.4 Against “Instrumental First” Replies

Here’s one potential line of response that a person might offer on Burge’s behalf. Someone might say that the important thing to focus on is instrumental rationality. That is to say, when we think about what it is to be a rational person, it is to be a person who does things to achieve their projects. Further, this person might say that it’s only on the basis of our practical agency that powers of representational agency, such as the merely impersonal and truly transpersonal aspects of reason, are attributed. This suggests that a two tiered approach is required. The instrumental aspects are to be promoted first and the impersonal and transpersonal aspects are to be promoted under what restrictions arise from promoting the instrumental aspects. I call this line of response “Instrumentality First”.

Here are some things that speak in favor of this line of response. First, it would allow Burge to accept MDO2. This revision would allow us to preserve the on-balance generic rationality of Tabitha’s and Agatha’s lies in their respective situations. After all, they’re putting the practical aspect of reasoning first, and the practical aspect of reasoning enjoys priority over the other aspects.

Second, it would be congenial, in at least some respects, to Burge’s own remarks about the fixation of representational functioning.<sup>4</sup> With respect to the capacity for perceptual representation, Burge claims that the attribution of perceptual representational capacities is driven by the demands of successful behavioral explanations. Not all agents are attributed representational capacities; they exhibit action without the need for mediating representations. However, when representational capacities are attributed in behavioral explanation, the specific representational capacities attributed are restricted by the practical situation of the animal; i.e., its characteristic environment in which the

---

<sup>4</sup>Burge (2009b); Burge (2010), pp. 319-342

behavior takes place as well as the needs and limitations of animal itself.

Despite these attractions, I think that Burge cannot accept this response. What's being proposed here is a substantive revision to the unity requirement according to which unity is restricted by considerations of explanatory priority. This revision allows for MDO2. But it also fails to constitute a rejection of MDO1. It fails as rejection of MDO1 because it renders MDO1 trivially true. It does this because it would entail that the antecedent of MDO1 is false.

The antecedent of MDO1 would be rendered false because the functional unity argument would be invalid. On the proposed revision to the unity requirement, anytime someone promotes the instrumental aspect at the expense of the merely impersonal and truly transpersonal aspects of reason, they're automatically unified. Accordingly, a person who promotes the instrumental aspect of reason at the expense of the impersonal aspects of reason is not disunified and is hence not irrational. This is problematic for Burge's functional unity argument in two ways. First, it would mean that the argument doesn't work. Second, it would mean that we must accept both premises of the Murderer-at the door objection. On these grounds, I recommend that Burge does not endorse this reply.

## 6.5 Against “Korsgaardian Unity” Replies

An alternative route to accepting MDO2 and rejecting MDO1 is to take a cue from another contemporary neo-Kantian, Christine Korsgaard. In her (1996) and her (2009), Korsgaard offers an account of personhood and agency according to which the function of a rational agent is to “pull themselves together”. The idea here is that, as people go through life, they find themselves cast in what seems to be a broad mishmash of roles, identities, and contexts. An irrational person, in this case, would be the “wanton”. The

“wanton”, when confronted with the disparate mishmash, denies their agency and merely allows the ebb and flow of the situation dictate the composition and priorities of their disparate roles and identities.

In contrast to the “wanton”, a successful rational agent is one who unifies their disparate roles and identities under a single “constitution”. In doing so, they *constitute* themselves as rational agents and people. They also promote their agency by making themselves active, causally relevant forces in the world.

There’s something to be said for this proposal. As I understand it, this proposal brings to bear an alternative account of the unity requirement. This Korsgaardian unity requirement is that agents themselves have to unify their functions under a constitution. A constitution, to a first approximation, is a description that, for the agent, settles questions of what to do and who they are.

This is certainly the sort of proposal that would allow Burge to hold on to MDO2. The explanation would be that Tabitha and Agatha would each be rational because they made a decision about the kind of person they wanted to be. They would each be pulling themselves together under the identity of being a person who protects people from imminent danger .

Furthermore, Korsgaard offers the Korsgaardian unity requirement, as I call it, as part of a broader Kantian meta-ethical program. This is tentative reason for cautious optimism. Korsgaard is looking to develop the categorical imperative in ways that address its faults. I’m willing to grant, for the sake of discussion, that Korsgaardian unity helps enough to differentiate the rationality of lying in murder at the door cases from other cases of lying. I’m even willing to grant that Korsgaardian unity is enough to support the validity and soundness of an appropriately limited version of the functional unity argument. Such a limited version would make an exception for murderer at the door cases.

I am not, however, willing to grant that Korsgaardian unity is an acceptable substitute for Burge's standard unity requirement. What's problematic about Korsgaardian unity, even after we grant all the things I'm willing to grant to Korsgaard, is that Korsgaardian unity seems to result in a hyper-intellectualized account of rational agency. Let me explain.

As I have mentioned elsewhere, a hyper intellectualized account of a mental capacity, such as an account of rational agency, is one that ties the possession of mental capacities to the possession of other, problematically demanding mental capacities. For example, in the philosophy of rational agency, Donald Davidson's account of rational agency is hyper-intellectualized because it requires creatures to be radical interpreters in order to count as agents. This is implausibly stringent because it seems as though there are a wide range of agencies, including mentally endowed agents, who fall short of being Davidsonian radical interpreters. For example, it seems implausible that people living with autism are all able to satisfy Davidson's requirement. However, people living with autism do possess minds and rational agency.<sup>5</sup> Hence, Davidson's account fails because it is hyper-intellectualized.

I think that Korsgaard's alternative unity requirement is hyper-intellectualized because, as she seemingly deploys it, the requirement conflates an important distinction between generic and critical rationality. Generic rationality is the set of standards associated with generic reason, as described in what I have called Burge's Veridicality account of rationality. Critical rationality is the set of standards associated with critical reasoning. Critical reasoning is the capacity to reason about how to reason. Critical reasoning involves the ability to meet some baseline standards for veridically representing reasons *as reasons*. It also involves the ability to meet some baseline standards for veridically representing and concepts of properties of reasons, such as concepts warrant, justifica-

---

<sup>5</sup>Andrews (2002) elaborates on this objection in convincing detail.

tion, reliability, logical inference, deductive validity, inductive strength, rationality, moral permissibility, and so on.

Looking back on the example of Davidson's hyper-intellectualization of agents, we might say that his requirements on agency conflate generic and critical reasoning. The ability to veridically radically interpret a behavior as a person's linguistically significant action requires, on Davidson's view, a whole host of critical reasoning capacities. These capacities include the abilities to articulate a theory of truth and a theory of reasons.

Now, we might say that Korsgaardian unity makes sense as a requirement of critical reasoning. It seems right to say that "pulling yourself together" and "uniting yourself under a constitution" involves scrutinizing your reasons as reasons, scrutinizing your identities as identities, and then imposing upon them an order of satisfaction that's sensitive to their rational properties. This seems like critical reasoning because Korsgaardian unity essentially involves reasoning about reasons.

Although Korsgaardian unity makes sense as a requirement of critical reasoning, it doesn't make sense as a requirement of generic reasoning and generic rationality. Again, it seems onerous to require that all rational agents must achieve the critical reasoning in order to be able to form any reasons, i.e., beliefs, desires, and plans, in the first place. Korsgaardian unity, offered as an account of generic reasoning, seemingly requires that a person be able to assess their reasons as reasons before they are able to reason about first order, worldly events. For example, it seemingly requires that a person be able to conceive of reasons as reasons in order to be able to form any reasons at all. Hence, I believe that Korsgaardian unity, as Korsgaard deploys it, is simply not going to help Burge rescue the functional unity argument from the Murderer at the Door objection. Korsgaardian unity is not ultimately helpful in rescuing the functional unity argument from the Murderer-at-the-Door objection because its deployment results in a hyperintellectualized account of agency.

## 6.6 Against “Goldilocks Critical Reasoning” Replies

The “Instrumental First” reply was too permissive. It was too permissive in the sense that it seemingly countenances as unified cases where a person is promoting the practical aspect of rationality at the expense of the impersonal aspects of rationality. The “Korsgaardian Unity” reply was too restrictive. It was too restrictive because it makes too demanding the possession conditions for rational agency. Intuitively, a person can be generically rational without being critically rational. Accordingly it is unrealistic to make critical rationality a necessary condition for generic rationality.

You might think that, a “Goldilocks”, “just right” reply would be to explicitly distinguish generic and critical reason, impose Korsgaardian unity as the standard for critical reasoning, and maintain that sometimes critical reasoning trumps generic reasoning. The idea would be to agree that it’s not true that it’s sometimes *generically* on-balance rational to assert lies rather than truths, but to maintain that it’s sometimes *critically* on-balance rational to assert lies rather than truths. Put another way, the functional unity argument is sound, but the exceptions we find in murderer at the door cases aren’t really cases of it being on-balance generically rational for a person to assert lies rather than truths. They’re instead cases of it being on-balance critically rational for a person to assert lies rather than truths. We’re accepting MDO1 while rejecting MDO2.

The motivation for this “Goldilocks Critical Reasoning” reply is the hope that explicit, but appropriately limited, appeal to critical reasoning can provide the desired special pleading for the morality and rationality of lying in murderer at the door cases. I think that critical reasoning abilities can be safely attributed to our lying protagonists in our murderer at the door cases. Critical reasoning abilities can be safely attributed because it helps frame and veridically explain the protagonists’ behaviors in our murderer at the door cases. This kind of reply would also have the benefits of the Korsgaardian

Unity reply without the problematic hyper-intellectualization, because the attribution of critical reasoning to our protagonists doesn't presume that critical reasoning abilities are required for generic reasoning. Instead, we can take our protagonists for what they are: rational agents who have reached a level of intellectual sophistication indicative of having achieved moral, and hence critical, rationality.

I'm nevertheless doubtful of the prospects for this reply. I'm doubtful because it leaves unclear how exactly generic and critical reasoning are may interact. We might think that critical reasoning is an ability that supercedes generic reasoning. This would mean that once a person has acquired the ability to critically reason, they no longer have to worry about promoting all three of the instrumental, impersonal theoretical, and transpersonal theoretical aspects of reason in their action. As critical reasoners, it is up to them to vindicate for themselves disunity among the functions of generic reason.

If critical reasoning supersedes generic reasoning, it seems like a new version of the rational deception objection emerges. We may call it, the "Critically Rational Deception Objection":

**CRDO1** It is sometimes on-balance critically rational for a person to assert lies rather than truths

**CRDO2** If it is sometimes on-balance critically rational for a person to assert lies rather than truths, then ambitious presumptivism is not true

**CRDO3** Therefore, ambitious presumptivism is not true

CRDO1 seems true in light of our intuitions about murderer at the door cases. As for CRDO2, it seems like we can import any one of the justifications a person might have offered in support of RDO2. After all, rational deception seemingly requires the ability to not just represent one's own goals, but it requires the ability to represent the audiences

reasons as reasons. If so, then rational deception requires the ability to critically reason. A proponent of the rational deception objection can simply insist that they were referring to critical reasoning all along when they were delivering their preferred justification of CRDO2.

Now, a person might reply that critical reasoning involves moral reasoning, and that moral reasoning impugns the morality, and hence critical rationality, of asserting lies rather than truths without special pleading. If so, we don't have to worry about CRDO2, because morality will do the job of reconciling CRDO1 with ambitious presumptivism. Hence, we can endorse a non-conciliatory response to with respect to the rational deception objection and a conciliatory response with respect to the critically rational deception objection.

I'm doubtful of the efficacy of this reply because I don't think that the attribution of critical reasoning is monolithic. I suspect that in order to be a critical reasoner, you only need some of the component aspects of critical reason, not all of it. For example, a manipulative sociopath can be a critical reasoner because they understand how their machinations impact the rational structure of other people's reasons. But, as sociopaths, they lack the moral aspects of critical reason.

Here is an additional complaint with the Korsgaardian unity requirement. Once we restrict the Korsgaardian unity requirement to critical reasoning, and not all reasoning, we take it out of the philosophical context where we could rest assured of the Korsgaardian unity requirement's moral philosophical promises. All bets are off. The angel Lucifer, of the poem "Paradise Lost", is surely a critical reasoner if ever there were one. As depicted in the poem, Lucifer adopted as a constitution for himself that "Evil be his good". I understand this as Lucifer reflecting on his ends and identities and uniting them under the maxim of pursuing the defeat of morality.<sup>6</sup> Since the adoption of a constitution is

---

<sup>6</sup>Although her discussion predates Korsgaard's, Anscombe is surely right about it being *prima facie*



enough for critical rationality, and since constitutions can be adopted for immoral ends, we cannot rest assured that Korsgaard-unified critical reasoners must be moral.

A different proposal about the relation between generic and critical rationality might go as follows. The relation between generic and critical rationality is not like the relation between two distinct species. Instead, it is akin to a relation between two members of the same species in distinct stages of development. Understood this way, there has only ever been a single unity requirement, and it's the original one that Burge offered. Critical rationality is achieved when a person acquires extra-generic aspects of reason, such as the aspects for logical reason and aspects for moral reasoning. But these acquisitions are on a par with the original instrumental, impersonally theoretical, and transpersonally theoretical aspects of reason. If so, critical reason doesn't shift agents into a different camp. Instead, it means that critically rational agents are beholden to more onerous rational requirements than those who are merely generically rational.

Although I think this alternative route seems more true to Burge's own conception of critical reason, it doesn't help Burge address the Murderers-at-the-Door Objection. It doesn't do so because it simply brings us back to where we started. We started in a place where the functional unity argument entailed that it's not rational for protagonists in the murderer at the door cases to assert lies rather than truths in the service of a moral end. From the standpoint of the original unity requirement on generic reason, the addition of aspects of critical reasoning don't suddenly make it okay to promote the instrumental and moral aspects at the expense of the transpersonally theoretical aspect. That would just be an occasion of disunity in a person's reasons. With this, I propose we reject the "Goldilocks Critical Reasoning" reply.

---

on-balance rational for Lucifer to will this way. See Anscombe (1957).

## 6.7 The “Deviant Contexts” Reply, Part One: Shiffrin on Suspended Contexts

Let’s consider one last class of replies on behalf of Burge against the Murderer-at-the-Door Objections. The essence of this reply would be to insist that murderer at the door cases constitutes a deviant cases where the functional unity argument, sound as it is, simply doesn’t apply. This means that when we affirm the soundness of the functional unity argument, we also discard Murderer-at-the-Door cases as outliers. Putting the point in the lingo of Burge’s Veridicality Account of Socially Rational Agency, we are simply insisting that Murderer-at-the-Door cases are among those situations where the Burge’s functional unity requirement is suspended. As such, the question of a person being unified across individual and social reasoning does not arise in such cases. Let’s categorize such replies as “Deviant Context” replies.

I think the best developed material for delivering a “Deviant Context” reply is Seana Shiffrin’s independently compelling account of “Suspended Contexts”. According to Shiffrin, a suspended context is one where the presumption of truthfulness has been suspended.<sup>7</sup> For Shiffrin, there are at least two kinds of suspended contexts, which she calls “epistemic suspended contexts” and “justified suspended contexts”. For present purposes, it will be helpful to re-label epistemic suspended contexts as ‘hearer suspended contexts’ and justified suspended contexts as ‘speaker suspended contexts’.

In a hearer suspended context (Shiffrin’s epistemically suspended context), facts about either the situation itself or the behavior of the participants “deprive the listener of the epistemic warrant to presume, in a predictive sense, that the speaker will tell the truth”.<sup>8</sup> For example, fictions involve misrepresentation, but the content of a fiction

---

<sup>7</sup>Shiffrin (2014), p. 16

<sup>8</sup>ibid.

might allow us to explore and enrich our conceptions of the world without the harms that may correlate from making the depicted scenarios real.<sup>9</sup> Similarly, matters of etiquette seem to serve essential aims where the content of the misrepresentation generates a speaker suspended context. A social interaction may be made all the more rewarding by a speaker's opening with the assertion of a compliment to the hearers. The self-esteem built by a successful complement would serve the goal of social inclusion. Arguably, inclusion, rather than presentation of truth, is the goal of the etiquette game. In these cases, the misrepresentation does not generate a speaker suspended context. Instead, these are all merely hearer suspended contexts. These are all contexts where a hearer does not have a reasonable expectation

A speaker suspended context (Shiffrin's so-called "justified suspended context") is a context where "the speaker's (potential) insincerity is reasonable and justifiable."<sup>10</sup> In a speaker suspended context, it should be accessible to all that insincerity is reasonable and justifiable.<sup>11</sup> As Shiffrin sees it, the crucial moral issue raised by murderer at the door cases, and lying in general, is in giving a satisfactory explanation of which contexts count as speaker suspended contexts.

Applied to murderer at the door cases, Shiffrin has this to say about the content centered generation of speaker suspended contexts:<sup>12</sup>

Given the compulsory ends morality supplies us, we would not reasonably use communication to further an evil end, and we cannot reasonably expect others to supply us with the reliable warrants necessary to do so... One in hot pursuit of an evil end should have no reasonable expectation that the world or other people cooperate with her evil enterprise. She has no entitlement to

---

<sup>9</sup>Shiffrin (2014), p. 32

<sup>10</sup>Shiffrin (2014), p. 16

<sup>11</sup>ibid.

<sup>12</sup>(Shiffrin (2014), pp. 33-34

those warrants about the world that would materially further those ends.

As I understand her, Shiffrin is telling us that the reason why murderer at the door cases count as speaker suspended contexts is because these are cases where, if the murderer thought about it, they would realize that their ends make it reasonable for speakers to be insincere. Specifically, the murderer's end of killing someone is such that, if the murderer were to think about it, they should realize that they have made it reasonable for people to obfuscate. Thus, the murderer's activities generate a speaker suspended context. Since the murderer's activities generate a speaker suspended context, the games of reasoned discourse are also suspended. After all, it is in the games of reasoned discourse that the expectation of speaking truth has its natural home. The immorality of the murderer's aim is one such that no one would reasonably want to play the game of reasoned discourse with them.

Here's how Shiffrin's account of suspended contexts might constitute a Deviant Context reply on Burge's behalf. The reply might be something like this. In inquiring into the location of their intended victim, the liar is inviting their audience into participating in a game of reasoned discourse. Should their audience accept the invitation then discursively socially rational agency would call upon all participants to unify reason's aspects both individually and socially. But the murderer's ends constitute an activity that people could not reasonably be expected to participate in. After all, the end of his activity is someone's murder and people have moral reasons against participating murdering others. Thus, in lying, people who are in roles like those of Tabitha and Agatha are not so much contravening the unity of reason. They are instead declining an invitation to contribute to an activity that they have reason to oppose. In this way, the context is deviant because the murderer's evil ends make it deviant. In this way, it can be the case that the functional unity argument is sound but that at the same

time, the liars in Murderer-at-the-Door scenarios are still rational in asserting lies rather than truth. The liars in Murderer-at-the-Door scenario are not disunified because their lie constitutes their declining to enter into a game of reasoned discourse. If Shiffrin's account of suspended contexts really gets us all this, it seems like we now have a *prima facie* reason to reject MDO1

## 6.8 The “Deviant Context” Reply, Part Two: Potential Resistance considered

Shiffrin's account of suspended contexts strikes me as a very promising way to implement a defense of Burge's Functional Unity Argument. Still, I think that a little more needs to be said to address a lingering issue. The lingering issue with relying on Shiffrin's route to some version of the “Deviant Context” reply is that it doesn't seem to cover all cases. I grant that Shiffrin's suspended context explanation clearly helps us with the **Lying Tabitha** and **Lying Agatha** cases. I think this is so because these are cases where Josephine, the would-be murderer, should know that that morality compels Josephine to cease her course of action, and that it compels Tabitha or Agatha decline her invitations to participate in her course of action.

Let us consider one last variation of a Murderer-at-the-Door case. This is a fictional case loosely inspired by Allen Wood's commentary Kant.<sup>13</sup> It is also very loosely inspired by current events:<sup>14</sup>

### LYING TAMARA

State lawmakers have outlawed any kind of abortion except for surgical abortions of pregnancies of less than 7 weeks. They have also ratified laws award-

<sup>13</sup>Wood (2007) pp. 240-58

<sup>14</sup>MacMammon (2021)

ing life altering bounties to ordinary citizens who accurately report on anyone involved in an abortion in violation of the ban. Mary is secretly recovering from an illegal surgical abortion. The only person who knows of Mary's pregnancy and surgical abortion is Tamara. Tamara secretly administered Mary's surgical abortion and is secretly harboring Mary during her recovery. Tina knocks on Tamara's door and Tamara answers. Tina asks Tamara if Mary is with her so that they (Mary and Tina) could talk. Tamara knows that asserting the truth invites a significant risk of Tina figuring out what has happened and reporting it. Uncertain of how Tina would handle the truth, Tamara lies and says she hasn't seen Mary recently. Tina accepts this and goes on about her day.

I think many people would say that it's both moral and rational for Tamara to lie.<sup>15</sup> Morality demands that you protect yourself and others from harm. And one such harm is getting wrongfully imprisoned for being involved with the abortion of a pregnancy. Given Tamara's relationship with Mary, and given her uncertainty about how Tina would take the truth, it seems the rational thing for Tamara to do is to assert a lie rather than speak the truth.<sup>16</sup> In this case, no assumptions are made as to whether the inquirer has malicious ends. However, the political context, tied with uncertainty about the inquirer's intentions, seems to be enough to contribute to the rationality of Tamara's lying.

On the face of it, the **Lying Tamara** case poses a problem for Burge's Functional

---

<sup>15</sup>In elaborating on my reasoning here, I'm simply going to deny that the anti-choice position is the moral position in this scenario. On this issue, I follow Judith Thomson's landmark (1971). The fetus's right to life, if it has one, doesn't seem to translate into a right to a person's body. I think this is true, even if, as argued in Marquis (1989), killing the fetus would deprive it of a "future like ours", and even if the case of aborting a six week old fetus doesn't bear any interesting analogies to cases that would normally justify killing a normal human adult.

<sup>16</sup>As Shiffrin reports it, Wood thinks that Kant was specifically interested in talking about how it's morally wrong for people to perjure themselves when brought into court by corrupt political officials (Shiffrin (2014), p. 28.). Accordingly, it isn't so much that Kant is opposed to any and all lying; but that lying in a legal context would be quite harmful for the liar, even if the lie is for a good cause (ibid.). I will set Wood's idiosyncratic view aside.

Unity Argument. It would pose a problem because in lying to Tina, Tamara is promoting certain aspects of reason over the impersonal aspect of reason. At the very least, she is promoting her practical welfare—after all, there’s a bounty program for catching people involved in illegal abortions—over the presentation of truth to other people.

And on the face of it, it would seem as though Shiffrin’s suspended context explanation could be deployed to Burge’s rescue. The reasoning might go something like this. The socio-political climate is such that it is prudent for Tamara to hide Mary’s pregnancy and her administering Mary’s abortion. After all, the bounty program for reporting on people who have had illegal abortions is so high as to be transformative. Given the facts of the case, it is hard to see why Tamara would take the risk of asserting the truth.

The lingering issue here is that there is still a contrast between this case and the **Lying Tabitha** and **Lying Agatha** cases. In the explanation of the **Lying Tabitha** and **Lying Agatha** cases, it was the immoral ends of the murderer that generated a speaker suspended context. That is to say, the manifest immorality of the murderer’s ends makes it prudent for the liars to decline the invitation by asserting a lie.

In contrast, Tina does not obviously have any immoral ends. Her manifest end is to speak with Mary. It is not clear that it is publicly accessible that Tina’s ends give Tamara reason to not tell the truth. This means that it is not clear that Tina has generated a speaker suspended context when asking after Mary’s whereabouts. After all, even though the political climate in the State has made it prudent for Mary to hide her pregnancy and abortion, and even though that climate has also made it prudent for Tamara to hide her involvement in with Mary’s abortion, it’s not as though Tina has reason to think that Mary might have had one. Accordingly, classifying the **Lying Tamara** case as an instance of a speaker suspended context seems extreme, if not inaccurate.

We might reply by saying that the political climate in the State is so extreme that it makes it that the State’s ends have co-opted ordinary citizens’s ends. The idea here is

that the state has incentivized things so that they have essentially co-opted Tina's ends. And so it is manifest to everyone that Tina is now a tool of the state. As a tool of the state, it should be publicly accessible that Tina has given anyone involved in an illegal abortion a reason to lie.

I don't think this works either. Here are two reasons. First, it seems too extreme. For all that has been said, Tina may be asking out of genuine concern for Mary's well-being rather than a self-interested snoop for an abortion bounty. For all that has been said, Tina, like Tamara, is willing to support people who have want to abort a pregnancy of more than six weeks. And, for all that has been said, Tina might be willing to support people who, like Tamara, aide others in acquiring abortions.

Second, even if it's quite likely that the climate is so extreme that State's have co-opted her ends, it is still not accessible to Tina that her being a potential tool of the State has given Tamara a reason to lie to her. After all, that inference requires Tina to know that Mary had an abortion that Tamara administered. Tina, by hypothesis, does not know this. She is not even in a position to know that her being a potential tool of the state generates a speaker suspended context.

A potential reply to this might be to bite the bullet. We might say that politically extreme climates often dissolve what were once reasonable expectations to answers about once innocent questions. But this response might go against the spirit of Shiffrin's account of suspended contexts. Consider how Shiffrin thinks the theory should handle evildoers in general:<sup>17</sup>

...a broad authorization to misrepresent threatens to isolate moral agents in good standing from those wrongdoers with whom they may have interests and to whom they may have a duty to try to establish relations on nonadversarial terms.

---

<sup>17</sup>Shiffrin, (2014), pp. 39-40



Shiffrin offers this as a response to why the suspended contexts generated by “evil seekers” should not be taken to be so thorough or absolute. The idea is that overextending the suspension of sincerity impedes upon the satisfaction of other moral demands, like the restoration of the moral community. In order for social progress to be made people will have to risk letting others in on the truth.

Let’s consider how this might apply to the **Lying Tamara** case. A community where women’s reproductive rights has been curtailed cannot begin to comprehend the restoration of those rights without admitting that there have been “illegal” exercises of those rights. Even though Tina’s ends have been co-opted by the broader community, she must, eventually, be let in on the truth of people like Mary and Tamara. Participation in games of reasoned discourse between people like Mary, Tamara, and Tina is a pre-condition for Mary, Tamara, and Tina’s community to reconsider and restore women’s reproductive rights.

Furthermore, Mary and Tamara are still part of the community. Although the community, in their case, fills the role of the “evildoer”, they might still want to have a relationship with their community. Relationships are complicated; even more so are the relationships between would-be pariah’s and their communities. Mary and Tamara’s completely isolating themselves from their community would undermine their identities as community members.

These considerations seem to suggest that even by Shiffrin’s lights, her account of suspended contexts cannot constitute a Deviant Context defense of Burge’s Functional Unity Argument.

## 6.9 The “Deviant Context” Reply: A Final Defense

The **Lying Tamara** case seems to suggest that there are Murderer-at-the-Door cases that are not really cases of suspended contexts. This is not necessarily a problem *per se* for Shiffrin’s proposal. For Shiffrin doesn’t think we should always have the same responses, even for Murderer-at-the-Door cases. But this is certainly a problem for using Shiffrin’s proposal to rescue Burge’s Functional Unity Argument from the Murderer-at-the-Door objection.

I don’t think that the problems posed by the **Lying Tamara** case are decisive. I agree that a significant part of the problem is using Shiffrin’s account of suspended contexts to do things it’s not really meant to do. Shiffrin’s account is meant to articulate what’s so problematic about classical versions of the Murderer-at-the-Door cases. It is not meant to solve them. But for my purposes, I need to solve them. So, I think that Shiffrin’s account needs some modifications.

Here is my proposed modification. In addition to Shiffrin’s hearer suspended contexts and speaker suspended contexts, I propose we add a third category. The third category is what I call “socially suspended contexts”. A socially suspended context is one where a significant breakdown in the community gives the speaker prudential reason to obfuscate and dissemble.

There are, I believe, historical cases of socially suspended contexts. Consider, for example, the societal breakdown in the United States that led to slavery-abolitionists to operate the famed Underground Railroad. In order to maintain their economic power, lawmakers in slave states offered bounties for returning escaped slaves. Such a bounty system constitutes those state’s co-opting the agency of the citizenry into pursuing the state’s pro-slavery ends. As such, when a person not involved with the underground railroad comes across a conductor of the Underground Railroad, that person’s ends are

intermixed with the states, even if they do not realize it. And this means that the state should expect that they have given conductors of the Underground Railroad reason to not participate in a sincere game of reasoned discourse.

Let me be clear: my socially suspended contexts are really just a device for expanding Shiffrin's category of speaker generated contexts to cover cases where the community's ends have entangled themselves with individual citizen's ends. This is important because it covers cases that seem like they should be speaker suspended contexts, but for which it's not necessarily publicly accessible to everyone that the context would be a speaker suspended context. When the community makes strides to mix its ends with those of individual citizens, the community inserts itself into those citizen's lives. You cannot interact with those citizens without interacting, as it were, to a potential agent of the community's sovereign will.

When the community mixes its ends with those of the citizens, it can do so in ways that need not be obvious to all of those citizens. After all, the ramifications of social policy are often more far reaching and convoluted for an ordinary citizen to be expected to grasp. And so it would make sense that a liar's reason for lying in such situations is publicly accessible, but not necessarily to the person being lied to. The reason is accessible to the community, as a source of the sovereignty. But this reason can, and often will, go over the heads of individual citizens.

Let's return to the **Lying Tamara** case. I contend that the **Lying Tamara** case is an instance of a socially suspended context. In setting up a bounty system, the state has co-mingled its ends with the ends of ordinary citizens. In doing so, it has turned ordinary citizens into the state's avatars or conduits. Accordingly, the conversation between Tamara and Tina is not just a conversation between Tamara and Tina. It is also a conversation between Tamara and the state. And in the conversation between Tamara and the state, we have what is essentially a speaker suspended context. For

Tamara is part of the community from which the state derives its knowledge, and the community knows, in a distributed-communal sense of knowing, that its ends give Tamara a reason to lie. In lying, Tamara is declining an invitation to engage in a game of reasoned discourse with the community. Since the lie is not a move in a game of reasoned discourse but instead the rejection of an invitation into such a game of reasoned discourse, Tamara is not occasioning disunity in the aspects of social reason.

## 6.10 Conclusion

I submit that Burge's functional unity argument is susceptible to a version of the Murderer-at-the-Door objection that dogged Kant's explanations of the Categorical Imperative. I also contend that this objection can be defeated. I contend that it can be defeated by appeal to an enriched version of Shiffrin's account of suspended contexts.

Shiffrin's account of suspended contexts offers us the space to reinterpret speaker's apparent lie, and hence disunified reason, into the speaker's declining an invitation to engage in social reason. Social reasoning, as such, is not compulsory. People cannot be expected to answer an inquirer's questions merely because the inquirer asked. I enriched Shiffrin's account of suspended contexts to handle cases where the social or communal climate is indicative of how the community has occupied the role of the "murderer". In doing so, I have preserved space for a proponent of Burge's functional unity argument to explain the rationality of a person's lying to unwitting conduits of an unjust regime. I hope that this ethical-political result is philosophically helpful.

# Bibliography

- Andrews, K. (2002). Interpreting autism: A critique of davidson on thought and language. *Philosophical Psychology*, 15(3), 317–332. <https://doi.org/10.1080/09515080210000061111>
- Anscombe, G. E. M. (1957). *Intention*. Harvard University Press.
- Bach, K., & Harnish, R. M. (1979). *Linguistic communication and speech acts*. Cambridge: MIT Press.
- BonJour, L. (1985). *The structure of empirical knowledge*. Cambridge, MA, USA: Harvard University Press.
- Brandom, R. (2000). *Articulating reasons: An introduction to inferentialism*. Harvard University Press.
- Brandom, R. B. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Harvard University Press.
- Bratman, M. (1987). *Intention, plans, and practical reason*. Cambridge: Cambridge, MA: Harvard University Press.
- Burge, T. (1977). Belief *De Re*. *Journal of Philosophy*, 74(6), 338–362. <https://doi.org/10.2307/2025871>
- Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, 4(1).
- Burge, T. (1993). Content preservation. *Philosophical Review*, 102(4), 457–488. <https://doi.org/10.2307/1523046>
- Burge, T. (2000). Frege on apriority. In P. Boghossian & C. Peacocke (Eds.), *New essays on the a priori* (pp. 11–42). Clarendon Press Oxford.
- Burge, T. (2003). Perceptual entitlement. *Philosophy and Phenomenological Research*, 67(3), 503–48. <https://doi.org/10.1111/j.1933-1592.2003.tb00307.x>
- Burge, T. (2009a). Five theses on de re states and attitudes. In J. Almog & P. Leonardi (Eds.), *The philosophy of david kaplan* (pp. 246–324). Oxford University Press.
- Burge, T. (2009b). Primitive agency and natural norms. *Philosophy and Phenomenological Research*, 79(2), 251–278. <https://doi.org/10.1111/j.1933-1592.2009.00278.x>
- Burge, T. (2010). *Origins of objectivity*. Oxford University Press.
- Burge, T. (2013). Postscript: ‘content preservation’. In *Cognition through understanding: Self-knowledge, interlocution, reasoning, reflection: Philosophical essays, volume 3* (pp. 254–284). Oxford University Press.
- Carnap, R. (1967). *The logical structure of the world* (R. Carnap, Ed.). Open Court.

## BIBLIOGRAPHY

---

- Church, A. (1940). A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5(2), 56–68. <https://doi.org/10.2307/2266170>
- Coady, C. A. J. (1992). *Testimony: A philosophical study*. Oxford University Press.
- Conee, E., & Feldman, R. (2004). *Evidentialism: Essays in epistemology*. Oxford University Press.
- Craig, E. (1990). *Knowledge and the state of nature: An essay in conceptual synthesis*. Clarendon Press.
- Davidson, D. (1984). *Inquiries into truth and interpretation*. Oxford University Press.
- Davidson, D. (1986). A coherence theory of truth and knowledge. In E. LePore (Ed.), *Truth and interpretation. perspectives on the philosophy of donald davidson* (pp. 307–319). Blackwell.
- Davidson, D. (2001a). *Essays on actions and events: Philosophical essays volume 1*. Clarendon Press.
- Davidson, D. (2001b). *Subjective, intersubjective, objective: Philosophical essays volume 3*. Clarendon Press.
- Davidson, D. (2004). *Problems of rationality*. Oxford University Press.
- Davidson, D. (2005). *Truth and predication*. Harvard University Press.
- Dummett, M. (1959). Truth. *Proceedings of the Aristotelian Society*, 59(1), 141–62.
- Elizondo, E. S. (2013). Reason in its practical application. *Philosophers' Imprint*, 13, 1–17.
- Faulkner, P. (2011). *Knowledge on trust*. Oxford University Press.
- Fodor, J. A. (1975). *The language of thought*. Harvard University Press.
- Frege, G. (1892/1997). On *Sinn* and *Bedeutung* (1892). In M. Beaney (Ed.), *The Frege reader* (pp. 151–171). Blackwell.
- Fricker, E. (1994). Against gullibility. In A. Chakrabarti & B. K. Matilal (Eds.), *Knowing from words*. Kluwer Academic Publishers.
- Fricker, E. (2006). Martians and meetings: Against Burge's neo-kantian apriorism about testimony. *Philosophica*, 78.
- Gallistel, C. R., & King, A. P. (2011). *Memory and the computational brain: Why cognitive science will transform neuroscience* (Vol. 6). John Wiley & Sons.
- Glanzberg, M. (2003). Against truth-value gaps. In J. C. Beall (Ed.), *Liars and heaps* (pp. 151–94). Oxford University Press.
- Goldman, A. (1979). What is justified belief. In G. Pappas (Ed.), *Justification and knowledge* (pp. 1–25). Boston: D. Reidel.
- Graham, P. J. (2000). Conveying information. *Synthese*, 123(3), 365–392. <https://doi.org/10.1023/a:1005162716568>
- Graham, P. J. (2012). Epistemic entitlement. *Noûs*, 46(3), 449–482. <https://doi.org/10.1111/j.1468-0068.2010.00815.x>
- Graham, P. J. (2018). Sincerity and the reliability of testimony: Burge on the a priori basis of testimonial entitlement. In A. Stokke & E. Michaelson (Eds.), *Lying: Language, knowledge, ethics, politics* (pp. 85–112). Oxford University Press.

## BIBLIOGRAPHY

---

- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66(3), 377–388. <https://doi.org/10.2307/2182440>
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, 3: Speech acts* (pp. 41–58). New York: Elsevier.
- Grice, H. P. (1989). Logic and conversation. In *Studies in the way of words* (pp. 22–40). Harvard University Press.
- Hegel, G. W. F. (1977). *Phenomenology of spirit* (A. V. Miller & J. N. Findlay, Eds.). Clarendon Press.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar* (A. Kratzer, Ed.). Blackwell.
- Jeffrey, R. C. (1965). *The logic of decision*. New York, NY, USA: University of Chicago Press.
- Kant, I. (1785/2018). Groundwork for the metaphysics of morals. In A. W. Wood (Ed. & Trans.), *Groundwork for the metaphysics of morals: With an updated translation, introduction, and notes* (A. W. Wood, Ed. & Trans.; pp. 1–73). Yale University Press.
- Kaplan, D. (1968). Quantifying in. *Synthese*, 19(1-2), 178–214. <https://doi.org/10.1007/BF00568057>
- Kaplan, D. (1989). Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from kaplan* (pp. 481–563). Oxford University Press.
- Kojève, A. (1969). *Introduction to the reading of hegel: Lectures on the phenomenology of spirit* (R. Queneau, Ed.). Cornell University Press.
- Kölbel, M. (2011). Conversational score, assertion, and testimony. In J. Brown & H. Cappelen (Eds.), *Assertion: New philosophical essays* (pp. 49–77). Oxford University Press.
- Korsgaard, C. M. (1996). *Creating the kingdom of ends*. Cambridge University Press.
- Korsgaard, C. M. (2009). *Self-constitution: Agency, identity, and integrity*. Oxford University Press.
- Kripke, S. A. (1980). *Naming and necessity*. Harvard University Press.
- Lackey, J. (1999). Testimonial knowledge and transmission. *Philosophical Quarterly*, 49(197), 471–490. <https://doi.org/10.1111/1467-9213.00154>
- Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8(1), 339–359. <https://doi.org/10.1007/bf00258436>
- MacFarlane, J. (2003). Future contingents and relative truth. *Philosophical Quarterly*, 53(212), 321–336. <https://doi.org/10.1111/1467-9213.00315>
- MacMammon, S. (2021). What the texas abortion ban does – and what it means for other states. *National Public Radio*. Retrieved September 1, 2021, from <https://www.npr.org/2021/09/01/1033202132/texas-abortion-ban-what-happens-next?t=1655099314000>
- Marquis, D. (1989). Why abortion is immoral. *Journal of Philosophy*, 86(4), 183–202. <https://doi.org/10.2307/2026961>

## BIBLIOGRAPHY

---

- Millikan, R. G. (1984). *Language, thought, and other biological categories*. MIT Press.
- Mossio, M., Saborido, C., & Moreno, A. (2009). An organizational account of biological functions. *British Journal for the Philosophy of Science*, *60*(4), 813–841. <https://doi.org/10.1093/bjps/axp036>
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.
- Price, H. H. (1964). *Perception*. Methuen.
- Pryor, J. (2005). There is immediate justification. In M. Steup & E. Sosa (Eds.), *Contemporary debates in epistemology* (pp. 181–202). Blackwell.
- Putnam, H. (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science*, *7*, 131–193.
- Quine, W. V. O. (1960). *Word and object*. MIT Press.
- Ramsey, F. (1926/2011). Truth and probability. In A. Eagle (Ed.), *Philosophy of probability: Contemporary readings* (pp. 52–94). Routledge.
- Rawls, J. (1955). Two concepts of rules. *Philosophical Review*, *64*(1), 3–32. <https://doi.org/10.2307/2182230>
- Redding, P. (2020). Georg Wilhelm Friedrich Hegel. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2020). Metaphysics Research Lab, Stanford University.
- Rescorla, M. (2007). A linguistic reason for truthfulness. In D. Greimann & G. Siegart (Eds.), *Truth and speech acts: Studies in the philosophy of language* (pp. 250–279). Routledge.
- Rescorla, M. (2009a). Assertion and its constitutive norms. *Philosophy and Phenomenological Research*, *79*(1), 98–130. <https://doi.org/10.1111/j.1933-1592.2009.00268.x>
- Rescorla, M. (2009b). Epistemic and dialectical regress. *Australasian Journal of Philosophy*, *87*(1), 43–60. <https://doi.org/10.1080/00048400802215653>
- Rescorla, M. (2015). Bayesian perceptual psychology. In M. Matthen (Ed.), *The oxford handbook of the philosophy of perception* (pp. 694–716). Oxford University Press.
- Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, *5*(6), 1–69. <https://doi.org/10.3765/sp.5.6>
- Salmon, N. U. (1986). *Frege's puzzle*. Ridgeview.
- Savage, L. J. (1954). *The foundations of statistics*. Wiley Publications in Statistics.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.
- Sellars, W. S. (1956). Empiricism and the philosophy of mind. *Minnesota Studies in the Philosophy of Science*, *1*, 253–329.
- Shieber, J. (2015). *Testimony: A philosophical introduction*. Routledge.
- Shiffrin, S. V. (2014). *Speech matters: On lying, morality, and the law*. Princeton University Press.
- Simion, M. (2021). Testimonial contractarianism: A knowledge-first social epistemology. *Noûs*, *55*(4), 891–916. <https://doi.org/10.1111/nous.12337>



- Simion, M., & Kelp, C. (2020). How to be an anti-reductionist. *Synthese*, 197(7), 2849–2866. <https://doi.org/10.1007/s11229-018-1722-y>
- Stalnaker, R. (1978). Assertion. *Syntax and Semantics (New York Academic Press)*, 9, 315–332.
- Stalnaker, R. (1999). *Context and content: Essays on intentionality in speech and thought*. Oxford University Press.
- Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. Methuen.
- Thomson, J. J. (1971). A defense of abortion. *Philosophy and Public Affairs*, 1(1), 47–66.
- Vargas, M., & Yaffe, G. (2014). *Rational and social agency: The philosophy of michael bratman*. Oxford University Press.
- Walton, D., & Krabbe, E. C. W. (1995). *Commitment in dialogue: Basic concepts of interpersonal reasoning*. State University of New York Press.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press.
- Wittgenstein, L. J. J. (1953). *Philosophical investigations*. New York, NY, USA: Wiley-Blackwell.
- Wood, A. W. (2007). *Kantian ethics*. Cambridge University Press.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.
- Wright, L. (1973). Functions. *Philosophical Review*, 82(2), 139–168. <https://doi.org/10.2307/2183766>