

The Prudential Significance of Care

By

© 2021

Dong-yong Choi

M.A. Yonsei University 2013

B.A. Yonsei University 2011

Submitted to the graduate degree program in Philosophy and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair Person: Dale Dorsey, Ph.D.

Ben Eggleston, Ph.D.

Jason Raibley, Ph.D.

Erin Frykholm, Ph.D.

Ani Kokobobo, Ph.D.

Date Defended: 29 March 2021

The dissertation committee for Dong-yong Choi certifies that this is the approved version of the following dissertation:

The Prudential Significance of Care

Chair Person: Dale Dorsey, Ph.D.

Date Approved: 29 March 2021

ABSTRACT

Prudence is a domain of action-evaluation, specifically concerned with the agent herself, in particular, with the agent's own welfare. For instance, if an agent can benefit either herself or a stranger, then in terms of prudence it is rational that the agent brings out advantages for herself unless she receives larger advantages as a result of benefiting the stranger. A feature of agents is that agents can exist over the course of time. In other words, it could be the case that an agent existed, is existing, and will exist. In regard to this feature of agents, a theory of prudence must be able to explain whether an agent's welfare at all times is equally important. According to a widely endorsed thesis of prudence (i.e. the temporal neutrality thesis), an agent's welfare at every moment of life is equally important. An agent's prudential reason to perform an action does not have a different strength depending on which part of life receives advantages/disadvantages as a result of performing the action. If option X benefits an agent's far future as much as option Y benefits the agent's near future, then the agent's reasons to choose the options have the same strengths.

In this dissertation, I suggest an alternative to the temporal neutrality thesis. If an agent either has friendships with or is psychologically related to other people, then in normative domains (e.g. morality and etiquette) the fact that the agent cares for the others is significant in evaluating the agent's actions. Based on this idea, I suggest the care thesis of prudence. First, I point out that agents have friendships with themselves at other times because agents and themselves at other times care for one another for one another's sake, and they know this fact about mutual caring. Furthermore, agents are psychologically related to themselves at other times because agents have their current mental states due to their previous mental states, and agents will have certain mental

states due to their current mental states. Second, based on these claims of intrapersonal friendship and psychological relation, I contend that the fact that at time T1 an agent cares for herself at time T2 more than for herself at time T3 is important in evaluating whether the agent's actions at T1 are rational. If at T1 an agent cares for herself at T2 more than for herself at T3, then it could be rational that the agent benefits herself at T2, even if she can bring out larger benefits for herself at T3.

After suggesting the care thesis, I argue that the care thesis is a more accurate understanding of prudence than the temporal neutrality thesis. In terms of prudence, it is appropriate that an agent prefers strong pain in the past to mild suffering in the future. The temporal neutrality thesis fails in explaining this phenomenon because, according to the temporal neutrality thesis, an agent's welfare at all times is equally important. In contrast, the care thesis can justify an agent's preference for past suffering. It is appropriate that agents prefer pain in the past to suffering in the future because agents care for themselves in the future more than for themselves in the past. The care thesis has a strength in that the care thesis can justify preferences for past suffering. However, critics could say that the care thesis is a problematic view of prudence because, according to this thesis, it could be permissible that an agent enjoys small benefits in the near future at the cost of significant benefits in the far future. In particular, critics could say that enjoying small benefits in the near future at the expense of significant advantages in the far future is a typical example of irrationality. I show that this criticism does not defeat the care thesis because it is rational to enjoy small benefits in the near future even in the case where the agent loses significant benefits in the far future.

The structure of this dissertation is as follows: In chapter 1, I explain crucial concepts for this dissertation. In chapter 2, I suggest the care thesis of prudence, and in chapter 3 I show that

the care thesis has explanatory power over the temporal neutrality thesis. Finally, in chapter 4, I dismiss the objection that the care thesis is a wrong view of prudence because the care thesis allows an agent to enjoy small benefits in the near future at the cost of significant advantages in the far future.

ACKNOWLEDGMENT

Without the encouragement of many generous and talented people, I would not have come to study in graduate school and would not have completed this dissertation project. I am grateful to each of these people for their contributions to completing my philosophy Ph.D. program and dissertation project.

I am grateful to my teachers at Yonsei University, Seoul, where I was first exposed to philosophy and where I did my B.A. and M.A. Thanks to the teachers' guidance, I fell in love with contemporary analytic philosophy (especially, with moral philosophy). Furthermore, I also want to thank LG Yonam Foundation for having funded me through my philosophy master program. I could not have graduated the master program if I had not received funding from LG Yonam Foundation.

I owe a debt of gratitude to Fulbright Scholarship Program. At the last phase of master program, I was experiencing economic hardships, so I decided to stop pursuing an academic career. However, thanks to Fulbright, I could apply for Ph.D. programs and as a result of it I am still pursuing an academic career. I especially thank Fulbright Executive Director Jai-ok Shim for her supports.

I am grateful to my teachers at the University of Kansas. They have shaped me into the philosopher I am today. The teacher, who is most responsible for whatever philosophical skills I have may is Dale Dorsey. I took almost every class taught by Dale. In my first semester in graduate school, I took his first-year graduate tutorial class and decided him as my philosophical role model. Through the courses, Dale taught me how to read philosophy and how to write philosophical papers.

Without Dale's support, I might not have been able to complete Ph.D. program at the University of Kansas.

I also want to thank Ben Eggleston, Jason Raibley, and Erin Frykholm for serving as the members of my dissertation committee. Furthermore, I am grateful for KU philosophy department and graduate school for having offered me funding for my project presentations and summer research.

Early versions of chapter 3 were presented at Minnesota Philosophical Association Conference and Illinois Philosophical Association Conference. Furthermore, the drafts of chapter 1 and 2 were presented at American Philosophical Association Annual Meetings (2021 pacific and 2021 central), Wisconsin Philosophical Association Conference, and United Conference of Korean Philosophers. Many people provided outstanding feedbacks on those occasions. I thank all of these people.

I have benefitted tremendously from countless interactions with my friends in graduate school. I am extremely grateful to Ben Caplan, Eileen Nutting, Nadia Ruiz, Taylor Hunt, Bada Kim, Ross Gilmore, Jong-jun Jeon, and Chulmin Yoon. I owe a huge debt of gratitude to my friends.

I want to thank my parents, In-kwon Choi and Min-ja Park. Without their unconditional supports, I could not complete this Ph.D. program. 어머니 아버지 감사합니다 (thanks for your love).

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGMENT	vi
1. INTRODUCTION.....	1
1.1 Prudence.....	1
1.2 Welfare.....	3
1.3 Project	7
1.4 Meaning	11
1.5 Conclusion.....	13
2. AN ALTERNATIVE TO THE TEMPORAL NEUTRALITY THESIS	16
2.1 Introduction	16
2.2 The Temporal Neutrality Thesis	17
2.3 An Alternative	22
2.3.1 Intrapersonal Friendship	23
2.3.2 Psychological Relation	30
2.3.3 The Care Thesis.....	36
2.4 Conclusion.....	38
3. FUTURE BIAS.....	39
3.1 Introduction	39
3.2 The Surgery Case.....	40
3.3 The Beneficial Attitude Reply.....	43
3.4 The Irrational Regret Reply	45
3.5 The Weak No Regrets Reply	50
3.6 The Motivation Objection.....	55

3.7 Conclusion.....	60
4. NEAR-TERM BIAS	61
4.1 Introduction	61
4.2 The Near-Term Bias Objection	62
4.3 The Arbitrariness Argument.....	64
4.4 The Instrumental Irrationality Argument.....	66
4.5 The Foreseen Regret Argument.....	70
4.6 The Awry Agency Argument	73
4.7 The Extreme Bias Argument.....	78
4.8 Conclusion.....	82
5. CONCLUSION.....	84
BIBLIOGRAPHY.....	86

CHAPTER 1

INTRODUCTION

1.1 Prudence

Many of our decisions center around ourselves. For instance, when we choose our life projects, we consider whether those projects are beneficial for us. Even if certain life projects are conducive to making our societies better, we might not select the projects if the projects cause significant harms for us. Among various normative domain, prudence is an exclusively self-regarding domain of action-evaluation, so a correct understanding of this domain would be helpful when we make our daily decisions. In this dissertation, I will investigate the nature of prudence. In particular, I will argue that, if at time T1 an agent cares for herself at time T2 more than for herself at time T3, then in terms of prudence it could be permissible that at T1 the agent makes decisions to benefit herself at T2 rather than herself at T3. This is because the fact that at T1 the agent cares for herself at T2 more than for herself at T3 is the agent's prudential reason to perform an action for her T2-self. For instance, if an agent cares for herself in youth more than for herself in old age, then it could be permissible that the agent spends resources for her youth, even if she loses her chances to benefit her old age.

Before investigating the nature of prudence, in this chapter, I will explain crucial concepts which I use in this dissertation. Let me start this preliminary work with the concept of prudence. As mentioned above, prudence is a domain of action-evaluation specifically concerned with the agent herself, in particular, with the agent's own welfare. Roughly speaking, while morality deals with what we owe to each other, prudence addresses what we owe to ourselves. Suppose that agent A has two options. If agent A chooses the first option, then she can save agent B from hunger. If agent A selects the second option, then she can enjoy leisure activities. In the domain of morality, agent A is required to choose the first option because agent B is in a desperate situation. In contrast, the prudentially rational thing to do is that agent A enjoys leisure activities unless she receives larger benefits as a result of saving agent B. For instance, if the pleasure from the beneficial act is stronger than the pleasure from her leisure activities, then in terms of prudence agent A should help agent B. However, if the beneficial act is not conducive to obtaining larger advantages, then it is rational that agent A enjoys leisure activities, for the ultimate goal of prudence is to benefit an agent herself.

The example above shows that prudence evaluates the rationality of actions. However, besides actions, this self-regarding normative domain also evaluates which attitudes are appropriate. To put this another way, an agent has prudential reasons for her attitudes as well as for her actions. Let me show this point with an analogy from morality. Imagine that agent A is close friends with agent B. They have spent lots of time together since they were kids, and they have kept this relationship throughout their entire lives. Agent B is sick, and she will receive a painful surgery to cure her disease. In this scenario, the fact that the surgery is painful makes agent A's being distressed appropriate. Furthermore, the fact that the treatment will save agent B's life makes agent A's being relieved fitting to the situation where agent B undergoes a painful operation.

On the contrary, agent A's indifference to the painful operation is inappropriate. Agent A has an obligation to care about her friends. Therefore, agent A is not doing her obligation if she is indifferent to the fact that her friend (i.e. agent B) will receive a painful operation in order to cure a disease.

An agent has moral reasons for her attitudes as well as for her actions. In the same vein, it is reasonable to suppose that prudence also evaluates both actions and attitudes.¹ Imagine that an agent knows she will have to undergo one of two surgeries, but she is not told which surgery she is going to receive. The agent learns that the first surgery will involve 1 hour of pain, and the second operation will involve 10 hours of suffering. Regardless of which surgery the agent receives, the surgery will save the agent from a disease. In this scenario, it is appropriate that the agent is distressed because she will go through a painful surgery. Moreover, it is also fitting that the agent is relieved because the surgery will save her life from the disease. Besides these two attitudes, another attitude is appropriate in this situation. If the agent hopes to receive the first operation, then this attitude is prudentially appropriate. In other words, other things being equal, the agent should hope to undergo the first surgery because the first surgery is less painful to the second operation.

1.2 Welfare

In the discussion of prudence, the concept of welfare is crucial because prudence evaluates actions depending on whether the actions increase the agent's own welfare. Welfare is the notion to

¹ Guy Fletcher also endorses this position. See Fletcher, "Taking Prudence Seriously" in *Oxford Studies in Metaethics: Volume 14*, ed. Russ Shafer-Landau (Oxford: Oxford University Press, 2019), 70-94.

evaluate an agent's state. In particular, an agent's level of welfare indicates how well the agent fares at a temporal location. If agent A's level of welfare is high at a certain time, then it indicates that agent A is faring well at the time. On the contrary, if agent A's level of welfare is low at a temporal location, then it means that agent A is not faring well at the temporal location. In regard to the nature of welfare, there are two kinds of theories: Subjective theories of welfare and objective theories of welfare. Chris Heathwood accounts for these two kinds of welfare theories as follows:

[A] theory is subjective just in case it implies the following: that something is intrinsically good for someone just in case either (i) she has a certain pro-attitude toward it, or (ii) it itself involves a certain pro-attitude of hers toward something [...] Correspondingly, objectivism about well-being is the view that at least one fundamental, intrinsic human goods does not involve any pro-attitudes on the part of the subject.²

Desire fulfillment theory, perfectionism, objective list theory, and hedonism are four main theories of welfare. According to the above classification of welfare-theories, desire fulfillment theory belongs to the group of subjective theories. Perfectionism and objective list theory are objective theories of welfare, and depending on how the nature of pleasure is understood, hedonism is classified into either the group of subjective welfare theories or the group of objective welfare theories.

Desire fulfillment theory is a subjective theory of welfare in that, according to this theory, an item is intrinsically good for an agent if, and only if, the agent wants to have the item. If an agent does not want to have an item, then it is impossible for the item to increase the agent's

² Heathwood, "Subjective Theories of Well-being" in *The Cambridge Companion to Utilitarianism*, eds. Ben Eggleston and Dale E. Miller (New York: Cambridge University Press, 2014), 205.

welfare unless the item brings out other intrinsically valuable items. Within the framework of desire fulfillment theory, two factors decide an item's intrinsic prudential value. The first factor is how long an agent desires to obtain an item. If an agent wants to have an item for a long period of time, then the item's intrinsic prudential value is larger than if the agent wants to have the item for a short period of time. The second factor is the strengths of desires. If an agent's desire to obtain an item is strong, then the fact that the agent's desire is satisfied increases the agent's level of welfare more than if the agent's desire to obtain the item is weak. Similarly, the fact that an agent's desire is frustrated is the sole ground to evaluate that the agent is not faring well at a temporal location.³

Perfectionism is an objective theory of welfare because this theory assumes that if an agent exercises the agent's essential capacity, then the agent's level of welfare increases, even if the agent does not care about whether she exercises the capacity. In particular, the fact that an agent exercises her essential capacity increases the agent's level of welfare. According to perfectionism, each species has its own essential capacities. For instance, homo sapiens's essential capacities are theoretical and practical rationalities. This is why if a member of homo sapiens makes an achievement in mathematics, then the member's level of welfare increases. In the same vein, if an agent's essential capacities are deteriorated, then the agent's level of welfare decreases.⁴ Unlike perfectionism, objective list theory assumes that more than one item is intrinsically valuable for an agent independent of the agent's attitudes. For instance, an objective list theory might assume

³ For more accounts of desire fulfillment theory, see Mark C. Murphy, "The Simple Desire-Fulfillment Theory" in *Noûs* 33 (1999); and Chris Heathwood, "Desire-Fulfillment Theory" in *The Routledge Handbook of the Philosophy of Well-Being* (London: Routledge, 2016), 135-147.

⁴ Thomas Hurka provides a systematic account for perfectionism. See Hurka, *Perfectionism* (Oxford: Oxford University Press, 1993). For arguments to support perfectionism, see Gwen Bradford, "Problems for Perfectionism" in *Utilitas* 29 (2017).

that an agent's moral goodness, rational activity, and knowledge are intrinsically good for the agent.⁵

Hedonism assumes that nothing but pleasure is intrinsically valuable for an agent. If an agent experiences pleasure, then the agent's level of welfare increases due to the fact that the agent has pleasant experiences. Similarly, the fact that an agent experiences pain decreases the agent's level of welfare. Depending on how the nature of pleasure is understood, hedonism of welfare is classified into either the group of objective theories or the group of subjective theories. According to felt-quality theories of pleasure, a mental state is an instance of pleasure because the mental state has a certain quale. If a felt-quality theory of pleasure is assumed, then hedonism is an objective theory of welfare, for there could be the case where an agent does not take a positive attitude toward prudentially valuable goods (i.e. pleasant experiences). In contrast, if it is assumed that an agent's positive attitudes toward mental states are what make the mental states pleasant, then hedonism is a subjective theory of welfare. For, within this understanding of pleasure's nature, an agent takes a positive attitude toward every intrinsic prudential good (i.e. every pleasant experience).⁶ In this dissertation, all discussions are ecumenical in regard to the above four theories of welfare except the case where the prudential values of project and meaning are discussed.

⁵ For more accounts of objective list theories, see Fletcher, "A Fresh Start for the Objective-List Theory of Well-Being" in *Utilitas* 25 (2013).

⁶ Fred Feldman provides systematic defenses for hedonism. See Feldman, *Pleasure and the Good Life: Concerning the Nature, Varieties and Plausibility of Hedonism* (Oxford: Clarendon Press, 2004). Ben Bramble also provides arguments to show that hedonism is a plausible view of welfare. See Bramble, "A New Defense of Hedonism about Well-Being" in *Ergo: An Open Access Journal of Philosophy* 3 (2016). Heathwood explains an objective version of hedonism and a subjective version of hedonism. See Heathwood (2014) 208-211.

1.3 Project

In chapter 3.3, I assume that if an agent completes her project, then the agent's welfare increases, even if at the moment of completion the agent no longer endorses the project. Let me explain why I endorse this view of project. Suppose that, throughout her entire life, agent A has worked hard in order to conquer Mt. Everest. Agent A started climbing up Mt. Everest, and she can arrive on the top of the mountain if she goes one step further. However, right before conquering Mt. Everest, agent A goes down the mountain because she is no longer interested in mountain climbing. In this case, one could contend that agent A's project abandonment does not have any negative impact on agent A's life because at the moment of abandonment agent A does not endorse the project. I think this claim about agent A's project abandonment is implausible. The moment where agent A abandons the project is just one of the project-related moments in agent A's life. In particular, given that agent A endorses the project throughout her life, the above claim about the abandonment puts too much significance on the fact that at the moment of abandonment agent A does not endorse the project.

Even if agent A no longer endorses her project, agent A's project abandonment has negative impacts on agent A's life because she has pursued the project throughout her life. In particular, I think the project abandonment has negative impacts on the life because, due to the abandonment, agent A loses her chance to increase her level of welfare. If agent A completed the project, then the project completion would have benefited the period of life where she pursues the project. However, agent A abandons the project, and it makes her lose the chance to benefit the part of life where she pursues the project. This is why the project abandonment has negative impacts on agent A's life. This account for the case above has two implications in regard to project completion. The first implication is about benefits from a project completion. Even if an agent no

longer endorses her project, the project completion can bring out benefits for the agent herself. Due to this feature, agent A's life would have received benefits if agent A completed the project. The second implication is about a project completion's beneficiary. If an agent completes her project, then the project completion benefits the period of life where the agent works for the project. Otherwise, agent A's project abandonment would not have had negative impacts on agent A's entire life.

Desire fulfillment theory can accommodate my view of project completion (i.e. an agent's project completion benefits the agent even in the case where the agent no longer endorses the project). According to desire fulfillment theory, even if an agent is not interested in her previous project, the project's completion benefits her (especially, the period of life where she pursues the project) because the completion fulfills her desire in the past (i.e. the desire to complete the project). Moreover, perfectionism can also explain my view of project's prudential value. For instance, perfectionists might claim that, since an agent has exercised her rationality while she pursues her project, the project's completion benefits every part of life (including the agent's past) where she exercises her rationality. Finally, objective list theory can provide a rationale for my view of project completion. For instance, advocates of objective list theory could assume that the fact that an agent completes her project increases the agent's level of welfare regardless of whether the agent still endorses the project. In particular, a project completion benefits an agent's temporal selves (i.e. the agent herself in the past or present) if the temporal selves contribute to completing the project. Unlike these theories of welfare, hedonism encounters a problem in explaining the fact that an agent's project completion increases the agent's welfare, even if the agent no longer endorses the project. Assuming that an agent fares better just in case the agent feels pleasure, an agent's welfare might not increase if the agent no longer endorses her project. This is because the

project completion would not make her pleased. Therefore, in regard to project completion's prudential value, I think hedonism is limited in explaining why project completion is prudentially valuable.

Critics might contend that my understanding of project completion is implausible because this understanding implies that it is possible to increase an agent's past welfare. In particular, critics might contend that it is impossible to increase an agent's past welfare because no one can bring out new events in the past. For instance, no matter what people do, people cannot save their past selves from already occurred tragic accidents, such as hunger. This objection relies on the idea that an agent's level of past welfare is determined just by what happens in the past. However, unless advocates of this objection suggest convincing arguments, there is no reason to endorse this assumption. This is because it is not a conceptual fact that an agent's level of past welfare is determined just by what happens in the past. To put this another way, in evaluating an agent's level of welfare, there is no conceptual restriction that an agent's past welfare should be determined just by what occurs in the past.⁷ Especially, given that a project abandonment can have negative impacts on an agent's whole life (i.e. an agent's past, present, and future), it is reasonable to say that events in the present (e.g. project abandonments) are factors to evaluate how well an agent fared.

Critics might suggest another objection to my understanding of project completion. Critics might say that, in the case above, agent A's project abandonment has negative effects on agent A's whole life. However, the reason is not that agent A loses her chance to benefit some periods of life

⁷ Donald W. Bruckner and Duncan Purves endorse the idea that an agent's event in the present can decide the agent's welfare in the past. See Bruckner, "Present Desire Satisfaction and Past Well-Being" in *Australasian Journal of Philosophy* 91 (2013); and Purves, "Desire Satisfaction, Death, and Time" in *Canadian Journal of Philosophy* 47 (2017).

(especially, the periods of life where agent A works in order to conquer Mt. Everest). If agent A conquered Mt. Everest, then agent A's whole life would have atemporally received advantages. Agent A has lost her chance to atemporally benefit her whole life. This is why agent A's decision has negative effects on agent A herself.⁸ This objection does not successfully defeat my understanding of project. Imagine that some of a tribe's members complete their personal co-project. In this case, it is implausible to say that the project completion benefits the tribe-as-a-whole. Given that certain members, not the tribe-as-a-whole, have worked in order to complete the project, it is difficult to believe that the project completion brings out advantages for the tribe-as-a-whole. In regard to this case, a reasonable position is that the project benefits the project's participants. Similarly, in the case above where agent A has climbed up Mt. Everest, it is reasonable to say that the conquest benefits the period of life where agent A makes efforts for the project.

If my account of project completion is correct, then as I will claim in chapter 3.3 a future-biased agent could lose chances to maximize welfare in her life. Suppose that agent A is no longer interested in project X. In particular, she has new interests in project Y, and she can complete project Y just in case she abandons project X. In this case, if an agent's project completion can benefit the agent even in the case where she no longer endorses the project, then it is possible that project X benefits agent A more than project Y. To put this another way, it could be the case that project X benefits the period of life where agent A works for project X more than project Y brings out advantages for the period of life where agent A pursues project Y. However, even if project X brings out larger benefits for agent A than project Y, if agent A cares about what happens in the

⁸ Krister Bykvist suggests a similar understanding of project completion. See Bykvist, "Comments on Dennis Mckerlie's 'Rational Choice, Changes in Value over Time, and Well-Being'" in *Utilitas* 19 (2007).

future more than she cares about what occurs in the past, then agent A might pursue project Y rather than complete project X. This is because future-biased agent A might consider her future interest (i.e. her interest in project Y) more important than her past interest (i.e. her interest in project Z), or she cares about her welfare in the future more than she cares about her welfare in the past.

1.4 Meaning

In chapter 4.7, I assume that if an agent's sacrifice is meaningful in that the sacrifice brings out larger advantages for the agent herself, then the fact that the sacrifice is meaningful is a reason to evaluate that the agent's life is going well for the agent. Let me explain what I mean by "meaning" and why sacrifice's meaning is relevant to the level of welfare. In the domain of prudence, the fact that with the passage of time an agent receives larger prudential goods is significant in evaluating the agent's life.⁹ Suppose that agent B is a beloved athlete. However, his early fame spoils him, and after several misconducts agent B lives a miserable life. On the contrary, a doppelganger of agent B lives a miserable life when he is young. However, due to the miserable experiences, the doppelganger becomes a successful coach.¹⁰ In this case, even if agent B's prudential goods at each time is the mirror image of the doppelganger's prudential goods at each time (i.e. young agent B has the same amount of prudential goods to his old doppelganger), in terms of prudence the

⁹ For instance, David Velleman, Michael Slot, Larry Temkin, and Joshua Glasgow endorse this position. See Velleman, "Well-Being and Time" in *Pacific Philosophical Quarterly* 72 (1991); Slot, "Goods and Times" in *Goods and Virtues* (Oxford: Oxford University Press, 1983); Temkin, *Rethinking the Good* (Oxford: Oxford University Press, 2012), chapter 3; and Glasgow, "The Shape of a Life and the Value of Loss and Gain" in *Philosophical Studies* 162 (2013).

¹⁰ I borrow this example from Dale Dorsey. See Dorsey, "The Significance of a Life's Shape" in *Ethics* 125 (2015), 304-305.

doppelganger's life is better than agent B's life. To put this another way, if one can decide whether she will live either agent B's life or the doppelganger's life, then in terms of prudence it is rational to choose the upward life (i.e. the doppelganger's life) rather than the downward life (i.e. agent B's life).

The relational view of welfare provides an account for the case above. According to the relational view, how well an agent fares during a period of time is not fully determined by the amount of prudential goods (e.g. the experience of pleasure or the fulfillment of desire) at each moment of the period. Otherwise, agent B would have fared as well as agent B's doppelganger. To put this another way, the four main theories of welfare (i.e. hedonism, desire fulfillment theory, objective list theory, and perfectionism) are limited in explaining how well an agent fares during a period of time. This is because these theories predict an agent's welfare at a time depending on the amount of prudential goods at the time. Advocates of the relational view contend that an agent's welfare during a period of time is also determined by how the agent's events during the period is related to one another. In particular, if an agent's hardship is meaningful in that the hardship brings out prudential goods for the agent herself, then the fact that the hardship is meaningful contributes to increasing the agent's welfare during the period of time where the hardship causes advantages. In the case above, the doppelganger's hardships are meaningful in that due to the experiences the doppelganger could become a successful coach. This is why the doppelganger's life is prudentially better than agent B's life. The fact that the hardships are meaningful increases the doppelganger's welfare.¹¹

¹¹ For more accounts of the relational view, see Velleman (1991); and Antti Kauppinen, "Meaningfulness and Time" in *Philosophy and Philosophical Research* 84 (2012).

The relational view is a plausible understanding of welfare in that advocates of this view can explain the significance of a life's shape (i.e. the fact that with the passage of time an agent receives larger advantages is important in evaluating the agent's level of welfare).¹² Based on this view of welfare, in chapter 4.7, I will make the following two assumptions: The first assumption is about the prudential value of sacrifice. I will suppose that if agent A receives benefits at time T2 as a result of her sacrifice at time T1, then the sacrifice is meaningful because due to the sacrifice agent A obtains advantages. In particular, since the sacrifice is meaningful in that agent A receives benefits, agent A's level of welfare increases. The second assumption is about who obtains benefits from meaningful sacrifices. I will assume that if agent A receives benefits at T2 as a result of her sacrifice at T1, then the fact that the sacrifice is meaningful makes both agent A at T1 and T2 fare better.

1.5 Conclusion

In this chapter, I have explained crucial concepts for this dissertation. Using these concepts, in the remaining chapters, I will investigate the nature of prudence. In chapter 2, I will suggest an alternative to the temporal neutrality thesis. The temporal neutrality thesis is a widely endorsed view of prudence. According to this thesis, an agent's welfare at all times is equally important in deciding the prudential statuses of the agent's actions. In particular, an agent's welfare at all times is equally important because an agent is herself regardless of when the agent exists. I will show that an agent has friendships with herself at other times. Moreover, I will point out that in terms of psychology an agent is related to herself at every moment of life. Based on these facts of

¹² Dorsey argues that the relational view explains the significance of a life's shape better than other views. See Dorsey (2015).

intrapersonal friendship and psychological relation, I will suggest the care thesis, according to which if agent A cares for herself at time T1 more than she cares for herself at time T2, then in terms of prudence it is permissible that agent A prioritizes her benefits at T1 over her advantages at T2. This is because caring within friendship/psychological relation is significant in evaluating actions.

In chapter 3, I will argue that the care thesis has explanatory power over the temporal neutrality thesis. First, I will introduce Derek Parfit's surgery case. This case shows that it is prudentially appropriate for an agent to prefer the scenario where she already experienced pain to the scenario where she will soon have suffering, even if the past pain is stronger than the future suffering. Second, I will point out that the temporal neutrality thesis encounters a problem in explaining this phenomenon of prudence. Assuming that an agent's past welfare is as important as the agent's future welfare, it seems inappropriate that an agent prefers strong pain in the past to minor suffering in the future. Third, I will argue that the care thesis can explain the above phenomenon of pain. According to the care thesis, it is appropriate that agents prefer the scenarios where they already experienced pain to the scenarios where they will soon have suffering. This is because agents care for themselves in the future more than for themselves in the past. Fourth, I will dismiss a criticism that the care thesis fails in explaining the surgery case. Moreover, I will refute three arguments, according to which the surgery case does not threaten the temporal neutrality thesis, so one has no reason to endorse the care thesis rather than the temporal neutrality thesis.

In chapter 4, I will reply to a criticism of the care thesis. Near-term bias is an attitude to care about what happens in the near future more than about what occurs in the far future. In the literature of prudence, it is widely assumed that this temporal bias is prudentially irrational. For

instance, people think that this temporal bias is not permissible because a near-term biased agent would choose today's minor benefits, even if she can receive significant advantages a few days later. Critics might contend that the care thesis justifies near-term biases, so the care thesis is a wrong view of prudence. I agree with this criticism in that if an agent cares for her near-future self more than for her far-future self, then the care thesis justifies the agent's near-term bias. However, the fact that the care thesis justifies near-term biases does not threaten the thesis because near-term biases could be rational. First, I will introduce five arguments to show that an agent's near-term bias is irrational. Second, I will dismiss these arguments against near-term biases. Based on this discussion of near-term biases, I will conclude that the mere fact that the care thesis justifies an agent's near-term bias does not show that the care thesis is a wrong understanding of prudential rationality.

CHAPTER 2

AN ALTERNATIVE TO THE TEMPORAL NEUTRALITY THESIS

2.1 Introduction

In the domain of prudence, an agent's welfare at every moment of life is significant in evaluating the agent's decisions. To put this another way, if an agent can benefit herself, then the agent has prudential reason to perform the beneficial act regardless of which period of her life receives advantages. According to the temporal neutrality thesis, not only does an agent's welfare at all times have prudential significance but an agent's welfare at every moment of life is equally important in evaluating the agent's actions. For instance, if an agent can bring out benefits for either herself in the near future or herself in the far future, then the agent should decide her action's beneficiary depending on whether she can bring out larger benefits for herself in the near future or herself in the far future. In this chapter, I will suggest an alternative to the temporal neutrality thesis. Especially, I will argue that since an agent is related in certain ways to herself at other times, in terms of prudence it is permissible that an agent prioritizes her welfare at time T1 over her welfare at time T2 if the agent cares for herself at time T1 more than she cares for herself at time T2.

The structure of this chapter is as follows: In section 2.2, I explain the temporal neutrality thesis. In particular, I introduce the personal identity argument to support this thesis. In section 2.3, I suggest the care thesis of prudence. I argue that the fact that an agent cares for herself at a time more than for herself at other times is significant in evaluating the agent's actions. Furthermore, I explain why the personal identity argument does not successfully support the temporal neutrality thesis.

2.2 The Temporal Neutrality Thesis

A widely endorsed view of prudence is the temporal neutrality thesis. According to this thesis, an agent's welfare at all times is equally important in evaluating the agent's actions. For instance, if agent A decides to enjoy minor benefits in the near future rather than significant advantages in the far future, then agent A's decision is prudentially irrational. Since an agent's welfare at every moment of life is equally important, agent A should choose the most beneficial option (i.e. the option to benefit herself in the far future) regardless of when she receives advantages. Henry Sidgwick, a proponent of the temporal neutrality thesis, describes his understanding of prudence as follows:

For my feelings a year hence should be just as important to me as my feeling next minutes [...]

Indeed this equal and impartial concern for all parts of one's conscious life is perhaps the most prominent element in the common notion of *rational* – as opposed to the merely *impulsive* – pursuit of pleasure.¹³

¹³ Sidgwick, *The Methods of Ethics* (London: Macmillan, 7th ed., 1907), 124, fn. 100.

David Brink makes a similar contention that “[p]rudence demands that an agent act so as to promote her own overall good. It is usually understood to require an equal concern for all parts of her life.”¹⁴

The temporal neutrality thesis can be described with the concept of reason-for-action. In the domain of prudence, the fact that action X brings out advantages for agent A is agent A’s prudential reason to perform action X. In particular, how much advantage action X brings out is a factor that determines the prudential reason’s significance. If action X brings out significant advantages for agent A, then agent A’s prudential reason to do action X (i.e. action X benefits agent A) is stronger than if action X causes small benefits to agent A. Besides the amount of advantage, other factors could determine the strength of agent A’s reason to do action X. For instance, if action X is against agent A’s deeply held ideals, then agent A’s reason to do the action might have weaker strength than if action X is compatible with agent A’s ideals. However, according to the temporal neutrality thesis, which part of agent A’s life receives the benefits is irrelevant to the prudential reason’s strength. Agent A’s reason to perform action X has the same importance regardless of which part of her life receives advantages as a result of her performing action X.¹⁵

As mentioned in chapter 1, not only do agents have prudential reasons for their actions but agents also have prudential reasons for their attitudes. Moreover, according to the temporal neutrality thesis, the timing during which an agent receives advantages/disadvantages is not

¹⁴ Brink, “Prospects for Temporal Neutrality” in *The Oxford Handbook of Philosophy of Time*, ed. Craig Callender (Oxford: Oxford University Press, 2011), 354.

¹⁵ Tom Dougherty, Preston Greene and Meghan Sullivan, Joyce Trebilcot, and Sidgwick take this position of prudence. See Dougherty, “Future-Bias and Practical Reason” in *Philosopher’s Imprint* 15 (2015); Greene and Sullivan, “Against Time Bias” in *Ethics* 125 (2015); Trebilcot, “Aprudentialism” in *American Philosophical Quarterly* 11 (1974), 204; and Sidgwick (1907), 124.

important in deciding the rationality of the agent's actions. Therefore, the coherent position for this thesis is that the timing during which an agent receives advantages/disadvantages is not significant in evaluating whether the agent's attitudes are appropriate. Suppose that agent A is in one of two situations. If she is in the first situation, then she receives a painful surgery for an hour in the near future. If she is in the second situation, then she undergoes the same surgery for ten hours in the far future. In this case, according to the temporal neutrality thesis, it is appropriate that agent A prefers the scenario where she receives a surgery in the near future to the scenario where she undergoes an operation in the far future. This is because the near-future surgery causes smaller pain than the far-future operation, and an agent's welfare at all times is equally important. If agent A prefers the far-future surgery, then in the perspective of prudence agent A's preference is inappropriate.

Two phenomena endorse the temporal neutrality thesis. The first phenomenon is, in the literature of prudence, it is widely assumed that an agent should sacrifice advantages in the present if the sacrifice causes larger advantages in the future. Sidgwick describes this prudential belief as follows:

[T]he present *pleasure or happiness* is to be foregone with the view of obtaining greater pleasure or happiness hereafter; but the principle need not be restricted to a hedonistic application, it is equally applicable to any other interpretation of 'one's own good', in which good is conceived as a mathematical whole, of which the integrant parts are realized in different parts or moments of a life time.¹⁶

Imagine that if agent A uses drugs, then agent A has pleasant experiences right now, but she loses her health. In contrast, if agent A does not use drugs, then she loses her chance to have pleasant

¹⁶ Sidgwick (1907), 381.

experiences, but she can keep her health. In this case, people believe that it is prudentially rational for agent A to sacrifice pleasure in the present because the sacrifice brings out larger advantages in the future. If agent A enjoys present benefits, then in terms of prudence the agent's decision is problematic.

Though it is widely believed that an agent should sacrifice present benefits for larger future advantages, a justification is needed since according to this belief an agent should sacrifice herself at a time. The concept of compensation is referred for this belief. Brink describes this justification as follows:

In this case, we should be able to justify sacrifices made at one point in an agent's life for the sake of some other period. A traditional rationale appeals to *compensation*. Now-for-later sacrifice is rational, because the agent is compensated later for her earlier sacrifice [...] In the intrapersonal case, benefactor and beneficiary are the same person, so compensation is automatic.¹⁷

According to this justification, an agent should sacrifice benefits in the present for larger advantages in the future because the agent's future self is the agent herself just as the agent's present self is herself. In other words, since an agent is identical to herself in the future, an agent's sacrifice is compensated if the sacrifice brings out larger advantages for herself in the future. This is why it is rational that an agent sacrifices present benefits if the sacrifice brings out larger future advantages.

The justification above relies on the idea that agents have prudential reason to benefit themselves in the future because agents are identical to themselves in the future. In particular, whether an agent receives advantages in the present or future is not important because the agent's

¹⁷ Brink (2011), 360-361.

future self is the agent herself just as the agent's present self is herself. Based on this idea of prudence, advocates of the temporal neutrality thesis might support their understanding of prudence as follows: If an agent in the present is identical to another agent in the future, then the former has prudential reason to benefit the latter. This is because an agent has prudential reason to bring out advantages for the agent herself regardless of when she exists. Agents are themselves, not someone else, regardless of at which temporal locations they exist. Therefore, an agent's welfare at every moment of life is equally important in evaluating whether the agent's action is irrational.

The second phenomenon for the temporal neutrality thesis is that whereas the fact that an agent herself will receive advantages is the agent's prudential reason to perform an action, the fact that a stranger obtains advantages is not a prudential reason for an agent to do an action. To put this another way, an agent has prudential reason to perform an action if the action benefits the agent herself in the future; but, unless an agent receives advantages in return, the agent does not have prudential reason to benefit a stranger. Advocates of the temporal neutrality thesis might claim that this phenomenon endorses the idea that personal identity is the reason why an agent has prudential reason to benefit herself in the future. Since an agent is personally identical to herself in the future, the fact that an agent can benefit herself in the future is the agent's prudential reason to perform an action. On the contrary, since an agent is not personally identical to a stranger, the fact that a stranger receives advantages is not a prudential reason for an agent to perform an action. An agent is personally identical to herself regardless of when she exists. This is why in the domain

of prudence an agent's welfare in the past and future is as important as the agent's welfare in the present.¹⁸

The discussion above shows that personal identity is crucial in justifying the temporal neutrality thesis. Especially, in order to endorse this thesis, advocates can articulate an argument as follows: First, if agent A is personally identical to agent B, then agent A has prudential reason to benefit agent B. This is why, for instance, an agent has prudential reason to perform a beneficial act for herself in the future, but an agent does not have prudential reason to perform an action for a stranger. Second, an agent is personally identical to herself at every other temporal location. For instance, an agent's past and future selves are the agent herself just as an agent's present self is the agent herself. Given that personal identity is the reason for an agent to benefit herself at other times, and an agent is personally identical to herself at every other temporal location, it is reasonable to conclude that an agent's welfare at all times is equally important in evaluating actions. In this dissertation, I will name this argument for the temporal neutrality thesis the *personal identity argument*.

2.3 An Alternative

According to the personal identity argument, the fact that an agent is personally identical to herself at other times is a reason for the agent to benefit herself at the other times. In particular, this argument assumes that personal identity is the sole relation which is prudentially significant. This is why an agent's welfare at all times is equally important in evaluating actions. In this section, I

¹⁸ Thomas Nagel, Robert Nozick, and Brink endorse this rationale in order to support the temporal neutrality thesis of prudence. See Nagel, *The Possibility of Altruism* (Princeton: Princeton University Press, 1970), part 2; Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 32-33; and Brink (2011), 360-361.

will argue that besides the relation of personal identity an agent has other relations with herself in the past and future, and these relations are prudentially significant in evaluating actions. Therefore, it could be the case that an agent's welfare at a time is important more than the agent's welfare at other times.

2.3.1 Intrapersonal Friendship

Aristotle says friendship is mutual. In particular, Aristotle describes his view of friendship as follows:

But people say that we ought to wish good things to a friend for his own sake. People describe those who do wish good things in this way, when the wish is not reciprocated, as having goodwill. For goodwill is said to count as friendship only when it is reciprocated.

Perhaps we should add 'and when it does not go unrecognized', since many have goodwill towards people they have not seen, but suppose to be good or useful; and the same feeling may exist in the other direction. They appear, then, to have goodwill to each other, but how could anyone call them friends when they are unaware of their attitude to one another? So they must have goodwill to each other, wish good things to each other for one of the reasons given, and not be unaware of it.¹⁹

According to Aristotle, if agent A has a friendship with agent B, then agent A wishes good things to agent B, for agent B's sake, and agent B also takes the same attitude toward agent A. However, having good wishes is not enough to have friendships with other agents. In order to have a friendship with agent B, agent A must be aware of the fact that agent B wishes good things to agent A herself, and agent B should be also aware of the fact that agent A wishes good things to agent B

¹⁹ Aristotle, *Nicomachean Ethics*, ed. Roger Crisp (Cambridge: Cambridge University Press, 2005), 1155b-1156a.

herself. If agent A fulfills these two conditions with agent B, then agent A has a friendship with agent B.

I agree with Aristotle's understanding of friendship's nature. In particular, I think agent A and B have a friendship with one another if they fulfill the following two conditions: The first condition is that agent A cares for agent B, for agent B's sake, and agent B also has the same attitude toward agent A. An agent shows various attitudes to other agents if the agent cares for the other agents, for their own sake. An agent feels satisfied when her cared one obtains benefits, desires to help the cared one when the one needs help, and sees the fact of her cared one's benefits as a reason to do an action.²⁰ If agent A has a friendship with agent B, then they mutually show these attitudes. Just as Aristotle says, even if agent A and B mutually care, they are not friends if they are unaware of the fact that they care for one another. The second condition to have a friendship is agent A knows that agent B cares for her, and agent B is also aware of the fact that agent A cares for her. For instance, if agent A has a friendship with agent B, then agent A might think agent B will help her when she needs help, for agent A knows that agent B cares for her. If agent A and B fulfill these two conditions for a friendship, then agent A and B are friends with one another.

Niko Kolodny might disagree with my conception of friendship. Kolodny contends as follows:

²⁰ For more accounts of caring, see Agnieszka Jaworska, "Caring and Internality" in *Philosophy and Phenomenological Research* 74 (2007), 560; and Jeffrey Seidman, "Valuing and Caring" in *Theoria* 75 (2009), 285.

Finally, relationships are *historical*. Whether I stand in a relationship to someone at a given time depends on some fact about our pasts. Kevin is my friend only if there has been a historical pattern of attitudes and actions between us.²¹

Kolodny might say that, in order to have a friendship, agent A and B should mutually care and know this fact about mutual caring. However, mutual care and knowledge about it are not sufficient to have a friendship. Especially, a shared history between agent A and B is a necessary condition for agent A to have a friendship with agent B. To put this another way, unless agent A and B have cared for one another and have known this fact about mutual caring, agent A is not friends with agent B.

A shared history between friends makes their friendship prudentially valuable. In particular, if friends have a long history of mutual care, then their friendship is more valuable for them than if they have a short history of mutual care. However, I do not think a shared history is necessary to have a friendship. If agent A and B mutually care, and they know this fact about mutual caring, then I think they start a relationship so called a *friendship*. Imagine that agent A is friends with agent B. For the last ten years, they have cared for one another and have known this fact about mutual caring. In this case, if one says agent A and B start their friendship in the fifth year, then this claim about the friendship's starting point is arbitrary. Given that agent A and B have cared for one another throughout their interactions, there seems to be no significant difference between agent A and B's fifth year and their fourth year. Similarly, in terms of friendship, their fourth year has no significant difference from their third year. The reasonable position for the starting point of friendship is that agent A and B's friendship has begun when they started mutually caring and knowing this fact of mutual care. At the time when they start caring, they might not have noticed

²¹ Kolodny, "Love as Valuing a Relationship" in *Philosophical Review* 112 (2003), 148.

that they have just begun a friendship, but this moment is the moment when they have become friends. Therefore, I think Kolodny's objection does not successfully refute my understanding of friendship.

Bennett Helm also might say my understanding of friendship is problematic, for she thinks "never to share activity with someone and in this way to interact with him is not to have the kind of relationship with him that could be called friendship, even if you each care for the other for his sake."²² In other words, according to Helm, shared activities are necessary components to have a friendship. In the discussion of friendship, the fact that an agent has participated in activities with other agents is important for three reasons. First, shared activities are one of the main processes through which an agent has friendships with other agents. For instance, while agents engage in joint pursuits, they might show virtuous characters toward one another, and it might make them mutually care. Second, shared activities benefit a friendship's participants, and the fact that the interactions bring out advantages for the participants makes their friendship prudentially valuable. For example, if religious friends cooperate together in order to follow their religion's teachings, then their friendship is better than if they barely interact with one another. Third, the fact that agents participate in activities together is an indicator to show that they are friends. For example, if an agent has never engaged in joint pursuits with others though she has a significant number of chances to do so, then it is reasonable to assume that the agent does not have a friendship with the others.

Though shared activities are conducive to having a friendship, interactions within a friendship make the friendship prudentially valuable, and shared activities are an indicator to show

²² Helm, "Friendship" in *The Stanford Encyclopedia of Philosophy*, Fall Edition (2017), URL = <<https://plato.stanford.edu/archives/fall2017/entries/friendship/>>.

that an agent has a friendship with another agent, shared activities are not necessary components to have a friendship. Suppose that agent A and B have not participated in activities together because they live far from each other. However, they strongly care for one another, and they know that they mutually care. For instance, agent A desires to help agent B whenever agent B needs her help, agent A is pleased when agent B's life goes well, and agent A sees the facts of agent B's advantages as practical reasons to perform actions. Agent B also has the same attitudes toward agent A, and agent A and B know that they mutually have these attitudes. In this case, it is difficult to believe that agent A does not have a friendship with agent B just due to the fact that they could not participate in activities together. In other words, it is unreasonable to say that agent A and B do not have special statuses (i.e. the statuses of friends) to one another just due to the fact that they could not engage in joint pursuits. The friendship between agent A and B might not be as valuable as the friendship between two actively interacting agents, but agent A has a friendship with agent B.²³

Critics might still claim that my understanding of friendship is problematic. Critics might say shared history and shared activity are necessary components to have a friendship, so my understanding is a wrong conception of friendship. In this chapter, I will not argue further to explain why mutual care and knowledge about it are sufficient to have a friendship. If it is difficult to endorse the idea that these two conditions are sufficient to have a relationship named *friendship*, then other names are available for this relationship, such as *quasi-friendship* or *friendship**. Regardless of which name we choose in order to refer to this relationship, I think this relationship

²³ For more accounts of friendship, see Elizabeth Telfer, "Friendship" in *Proceeding of the Aristotelian Society* 71 (1971); Samuel Scheffler, "Relationships and Responsibilities" in *Philosophy and Public Affairs* 26 (1997); Diana Jeske, "Friendship, Virtue, and Impartiality" in *Philosophy and Phenomenological Research* 57 (1997); and Simon Keller, *Partiality* (Princeton: Princeton University Press, 2013).

has a similar nature to a friendship, for in order to start this relationship agents should care for one another, and they should know this fact about mutual caring. Furthermore, this relationship is valuable due to the feature which also makes a friendship valuable. As Seth Lazar points out, friendship is one of the noblest achievements in a life because, in this relationship, agents care about each other's welfare for its own sake.²⁴ Due to the same reason, quasi-friendships are valuable because agents in these relationships care for one another, for one another's sake. Therefore, it is reasonable to contend that quasi-friendships and friendships are *allied genera*. In the discussion below, I will proceed assuming that quasi-friendship is a version of friendship. In particular, I will assume that in normative domains quasi-friendships have similar significances to friendships.

Agent A has a friendship with agent B if agent A and B meet two conditions. First, agent A cares for agent B, for agent B's sake, and agent B also has the same attitude toward agent A. Second, agent A and B are aware of the fact that they care for one another. An agent can meet these two conditions for a friendship with herself at a different time, so I think it is reasonable to contend that an agent can have a friendship with her temporal self. First, an agent and her temporal selves can care for one another, for one another's sake. For instance, an agent in youth could want to be healthy when she gets old, and an agent in old age could wish that she could have avoided the unnecessary hardships when she was young. Second, it is possible that an agent and her temporal selves know that they mutually care. For example, if an agent cares for her past self, then the agent might reasonably believe that her future self would care for her (i.e. the future self's past

²⁴ Lazar, "The Justification of Associative Duties" in *Journal of Moral Philosophy* 13 (2016), 50.

self). Moreover, if an agent cares for her future self, then the self would know she was cared by the agent.

In the domain of morality, it is permissible that an agent cares about her friend's welfare more than about a stranger's welfare. For instance, if agent A is friends with agent B, then it could be permissible that agent A benefits agent B, even if agent A can bring out larger advantages for a stranger. In the same vein, if an agent has friendships with herself at certain times, then I think it is permissible that the agent cares about her welfare at the times more than about her welfare at other times. Before explaining this idea in detail, let me briefly explain how I understand the normative significance of friendship. I endorse the appropriate response view of value. According to this view, "[t]o say that x is valuable is to say that it has a number of properties to which a given response, recognizable as a form of valuing, is appropriate."²⁵ I think friendship is prudentially valuable because this relationship is a stable realization of altruism. In other words, it is permissible that an agent cares about her friend's welfare more than about a stranger's welfare because a partial attitude toward a friend is a proper response to the agent's own achievement (i.e. altruism in friendship). Similarly, if agent A cares for friend B more than for friend C, then it could be permissible that agent A cares about friend B's welfare more than about friend C's welfare because, due to this pattern of caring, the friendship with friend B could be valuable more than the friendship with friend C.

Suppose that agent A and herself at time T1 care for one another and know this fact about mutual caring, but agent A and herself at T2 are indifferent to each other. In this scenario, I think it is prudentially permissible that agent A cares about her T1-self's welfare more than about her

²⁵ Lazar (2016), 45.

T2-self's welfare. This is because agent A is friends with herself at T1, and a partial attitude is a proper response to the friendship with herself at T1. Similarly, even if agent A has friendships with both herself at T1 and T2, if agent A cares for herself at T1 more than for herself at T2, then it could be permissible that agent A cares about her T1-welfare more than she cares about her T2-welfare. This is because, due to the fact that agent A cares for herself at T1 more than she cares for herself at T2, agent A's friendship with the T1-self could be valuable more than her friendship with the T2-self. In summary, if an agent has a friendship with herself at a certain time, then it could be permissible that the agent cares about her welfare at the time more than about her welfare at other times because the agent's friendship with herself at the time justifies the agent's partial attitude.

2.3.2 Psychological Relation

While investigating the concept of practical rationality, Derek Parfit suggests the following scenario:

My body is fatally injured, as are the brains of my two brothers. My brain is divided, and each half is successfully transplanted into the body of one of my brothers. Each of the resulting people believe that he is me, seems to remember living my life, has my character, and is in every other way psychologically continuous with me. And he has a body that is very like mine.²⁶

Let us name the agent with a damaged body *agent X*, the agent with agent X's left-brain *agent Y*, and the agent with agent X's right-brain *agent Z*. Suppose that, in this case (hereafter referred as *My Division*), agent X should allocate her resources before she receives the brain operation. Agent X can bring out a larger amount of advantages for strangers in the future, but agent X gives her

²⁶ Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), 254-255.

resources to agent Y and Z. Especially, agent X makes this decision because she cares for agent Y and Z more than she cares for strangers in the future. In this situation, unless agent X can bring out an enormous amount of advantages for future strangers, in the perspective of morality agent X's decision seems permissible. Similarly, after the brain operation, if agent Y and Z care for agent X more than they care for other deceased strangers, then it seems morally permissible that agent Y and Z accomplish agent A's uncompleted project even in the case where they can complete other deceased people's projects. In the domain of morality, it is justifiable that an agent cares about her own welfare more than she cares about a stranger's welfare. However, given that agent X is not personally identical to either agent Y or Z, personal identity is not the reason that justifies agent X's decision for agent Y and Z. Agent X is not personally identical to both agent Y and Z because agent Y is not personally identical to agent Z. Furthermore, it is arbitrary to contend that agent X is personally identical to one of agent Y and Z because agent Y and Z are the same in every aspect.²⁷

Another possible account for the phenomenon above (i.e. it is permissible that agent X, Y, and Z care about one another's welfare more than about other people's welfare) appeals to the concept of child-parent relationship. According to this account, agent X has a child-parent relationship with agent Y and Z. Just as parents give birth to their children, agent X gives birth to agent Y and Z. In the domain of morality, parents and their children have reason to care about one another's welfare more than they care about other people's welfare.²⁸ This is why it is permissible that agent X, Y, and Z benefit one another rather than other agents. I also think that parents and

²⁷ For the detail accounts of agent X's survival, see Parfit (1984), 199-345.

²⁸ For the accounts of parent-child relationship's moral significance, see Simon Keller, "Four Theories of Filial Duty" in *The Philosophical Quarterly* 56 (2006); and David Archard, "The Obligations and Responsibilities of Parenthood" in *Procreation and Parenthood: The Ethics of Bearing and Rearing Children*, eds. David Archard and David Benatar (Oxford: Oxford University Press, 2010).

their children have reason to be partial toward one another, but the framework of parent-child relationship is not conducive to understanding the phenomenon above. Even if parents do not care for their children, parents have reason to be partial toward their children. In particular, parents are required to benefit their children if they can benefit strangers just as much as they can benefit their children. Similarly, regardless of whether children care for their parents, children are required to benefit their parents if they can bring out the same benefits for either strangers or their parents. However, in *My Division*, if agent X, Y, and Z do not care for one another, then they seem to have no reason to be partial toward one another. For instance, if agent X cares for agent Y and Z just as much as she cares for other people, then it seems permissible that agent X brings out advantages for other people, even if she can benefit the strangers just as much as she can benefit agent Y and Z.

I think agent X's decision is permissible because agent X is psychologically related to agent Y and Z, and agent X cares for agent Y and Z more than for strangers. The fact that an agent cares for other agents can justify the agent's partial attitudes if the agent is related in a certain way to the other agents. Let me show this point with an example. Imagine that agent B and C are strangers to agent A. Even though agent C needs agent A's help more than agent B, agent A benefits agent B because she cares for agent B more than for agent C. In particular, agent A cares for agent B just because agent B came from the region where she used to live. In this case, agent A's decision is morally problematic. Agent A should have benefited agent C because agent C needed her help more than agent B. However, if agent A and B have worked together for good causes, and this is why agent A strongly cares for agent B, then it could be permissible that agent A benefits agent B even in the case where agent A can bring out larger advantages for agent C. In particular, unless agent A can bring out an enormous amount of advantages for agent C, it is permissible that agent

A benefits agent B rather than agent C. These two scenarios show that an agent's caring can have moral significance if the agent is related in a certain way to her cared agent. This is why even though agent A, in both of the scenarios, cares for agent B more than for a stranger, agent A's decision to benefit agent B is justifiable only in the scenario where agent A has worked with agent B.²⁹

Besides the fact that agents have engaged in joint pursuits for good causes, other facts could make caring significant. In particular, I think if an agent is psychologically related to other agents, then the fact that the agent cares for the other agents justifies the agent's partial attitudes toward the other agents. This is why in *My Division* it is permissible that agent X, Y, and Z care about one another's welfare more than about other people's welfare. If agent X did not care for agent Y and Z, then agent Y and Z's welfare would have been as important as other people's welfare. However, by stipulation, agent X cares for agent Y and Z more than she cares for other agents. In particular, agent X is psychologically related to agent Y and Z. This is why the fact that agent X cares for agent Y and Z justifies agent X's partial attitudes toward agent Y and Z. In other words, due to psychological relation, agent X's caring justifies agent X's decision for agent Y and Z. Similarly, if agent Y and Z care for agent X, then it is also permissible that agent Y and Z care about agent X's welfare more than about a stranger's welfare because they are psychologically related to agent X.

As shown above, *My Division* is not explicable within the framework of parent-child relationship. Even if parents and their children do not care for one another, parents and their children should care about one another's welfare more than about other people's welfare. However,

²⁹ Thomas Hurka defends this view while discussing partial attitudes toward compatriots. See Hurka, "The Justification of National Partiality" in *Drawing Morals: Essays in Ethical Theories* (New York: Oxford University Press, 2011), 219-237.

agent X, Y, and Z do not have to prioritize one another's advantages if they do not care for one another. My account above (i.e. a care-based account) can accommodate this phenomenon of prioritizing. Agent A's psychological relation with agent B does not require agent A to care about agent B's welfare more than about a stranger's welfare. This relation just makes agent A's caring significant. In Jonathan Dancy's terminology, the fact that agent A is psychologically related to agent B is an *enabler*, and due to this enabler the fact that agent A cares for agent B becomes agent A's reason to care about agent B's welfare more than about a stranger's welfare.³⁰ This is why, even though agent X is psychologically related to agent Y and Z, agent X does not have reason to prioritize agent Y and Z's benefits if agent X does not care for them. As this discussion shows, my account can explain *My Division*. Furthermore, this account does not have the other account's problem (i.e. the problem that a parent-child relationship approach encounters). Therefore, it is reasonable to say that the fact that agent A cares for agent B justifies agent A's partial attitude toward agent B.

My Division endorses the idea that if agent A is psychologically related to agent B, and agent A cares for agent B more than for agent C, then it is permissible that agent A cares about agent B's welfare more than she cares about agent C's welfare. This idea of caring is plausible because it can also make a feasible verdict in another case. In terms of morality, it is permissible that an agent cares about her teacher's welfare more than about a stranger's welfare. Similarly, it is also justifiable that an agent considers her student's welfare more important than a stranger's welfare. If one accepts my view of caring (i.e. caring within psychological relation is important), then one can explain these phenomena of partial attitudes toward teachers and students. As a result

³⁰ For detail accounts for the concept of enabler, see Dancy, *Ethics without Principles* (Oxford: Oxford University Press, 2004), chapter 5.

of interactions in class, teachers are psychologically related to their students. This is why if an agent cares for her teachers more than strangers, then it could be permissible that the agent benefits her teachers even in the case where she can bring out a larger amount of advantages for strangers. In the same vein, if an agent cares for her student, then in terms of morality it is permissible that the agent benefits her student, even if the agent can bring out a larger amount of advantages for a stranger.

I have shown that if agent A is psychologically related to agent B, but she has no relation to agent C, then the fact that agent A cares for agent B more than for agent C justifies agent A's partial attitude toward agent B. However, it does not have to be the case that agent A is psychologically related just to agent B for her partial attitude to be permissible. Even if agent A's mental states are equally related to agent B and C's mental states, if agent A cares for agent B more than she cares for agent C, then the fact that agent A cares for agent B more than for agent C can justify agent A's partial attitude toward agent B. Imagine that, in terms of psychology, agent A is equally related to her teacher and student. In this case, if agent A cares for her teacher more than for her student, and the reason is that agent A puts importance on the fact that she has learnt knowledges from her teacher, then it is permissible that agent A cares about her teacher's welfare more than about her student's welfare. On the other hand, if agent A cares for her student more than for her teacher because she considers it important that she has made impacts on the student's beliefs, then it is permissible that agent A cares about her student's welfare more than about her teacher's.

Caring matters in the domain of guest-serving etiquette as well as in the domain of morality. Imagine that agent A invites three guests to her place. In particular, one of the guests (i.e. agent B) is the one who has made significant influences on agent A's beliefs, desires, and intentions. Agent

A has two options. If she chooses the first option, then she can serve good deserts for her guests. In contrast, if she selects the second option, then she can serve an excellent desert for agent B, but the other guests receive decent deserts. In this case, it is appropriate in the sense of guest-serving etiquette that agent A selects the first option. However, if agent A cares for agent B more than for the other guests, then it seems also appropriate that agent A chooses the second option rather than the first option. In particular, I think it is appropriate that agent A chooses the second option because agent A is psychologically related to agent B, and agent A cares for agent B more than she cares for the other guests. This example of guest-serving etiquette and the above discussion of morality endorse the idea that, in normative domains, caring justifies partial attitudes if a caring agent is psychologically related to her cared ones. In the same vein, I think it is reasonable to contend that if an agent is psychologically related to herself at a certain time, then the fact that the agent cares for herself at the time justifies the agent's partial attitude toward herself at the temporal location.

2.3.3 The Care Thesis

In section 2.3.1, I argued that an agent can have intrapersonal friendships with herself at other times. In fact, an agent not only can have an intrapersonal friendship with herself at a different time but in most cases an agent does have a relationship with herself in the past and future. First, people care for themselves in the past and future as well as themselves in the present. For instance, people wish that they could have avoided the unnecessary pains, and people also want to enjoy high levels of welfare in the future. Second, people know that their past and future selves care for them just as they care for the temporal selves. For instance, people reasonably believe that they would not have made the decisions if they had known that the decision would sabotage their welfare, and people pursue projects because they have justified beliefs that they would complete

their projects if they can do so. Therefore, other things being equal, it is prudentially permissible that agent A cares about her welfare at time T1 more than about her welfare at time T2 if agent A cares for herself at T1 more than she cares for herself at T2. This is because, in this case, agent A's friendship with herself at T1 would be prudentially more valuable than her friendship with herself at T2.

I also argued that if an agent is psychologically related to herself at time T1, then the fact that the agent cares for herself at T1 justifies her partial attitude toward herself at T1. In almost every case, an agent is psychologically related to herself at every temporal location. For instance, an agent has certain beliefs, desires, and intentions because she had those mental states before. Even if an agent's mental states are different from her mental states in the past, the past mental states determined what kinds of mental states the agent would have. Similarly, an agent is psychologically related to herself in the future, so it is reasonable to contend that the fact that an agent cares for herself at a certain time justifies the agent's partial attitude toward herself at the time. In summary, caring makes intrapersonal friendship valuable because it realizes intimate reciprocal care. Furthermore, caring within psychological relation makes it permissible that an agent prioritizes her card one's welfare. Therefore, if an agent cares for herself at time T1 more than she cares for herself at time T2, then in evaluating actions and attitudes the agent's welfare at T1 is significant more than the agent's welfare at T2. I will name this view of prudence the *care thesis*.

In section 2.1, I introduced the personal identity argument for the temporal neutrality thesis. According to this argument, an agent's welfare at all times is equally important because an agent is herself regardless of when she exists. It is true that, in the domain of prudence, personal identity is what makes an agent's welfare in the future important to the agent herself in the present. This is

why an agent does not have prudential reason to benefit a stranger unless she receives advantages in return. However, the fact that personal identity matters does not justify the temporal neutrality thesis because not only does an agent have the relation of identity with herself at other times but an agent also has psychological relations and friendships with herself in the past and future. In particular, due to these relations, if an agent cares for herself at T1 more than for herself at T2, it is permissible that the agent cares about her T1-welfare more than about her T2-welfare. Therefore, the personal identity argument fails in showing that the temporal neutrality thesis is a right view of prudence.

2.4 Conclusion

In this chapter, I suggested a new view of prudence: The care thesis. According to the care thesis, if agent A cares for herself at time T1 more than for herself at time T2, then it is permissible that agent A cares about her welfare at T1 more than about her welfare at T2. This is because, due to this pattern of caring, agent A's friendship with her T1-self is valuable more than her friendship with her T2-self. Furthermore, since an agent is psychologically related to herself at other times, the fact that agent A cares for herself at T1 more than for herself at T2 justifies agent A's partial attitude toward herself at T2. The care thesis is incompatible with the temporal neutrality thesis, so it should be examined which thesis is a more accurate understanding of the self-regarding domain.

CHAPTER 3

FUTURE BIAS

3.1 Introduction

In chapter 2, I suggested the care thesis of prudence. According to the care thesis, if an agent cares for herself at time T1 more than for herself at time T2, then it is permissible that the agent cares about her welfare at T1 more than about her welfare at T2. This view of prudence is incompatible with a widely endorsed view of prudence: The temporal neutrality thesis. Unlike the care thesis, the temporal neutrality thesis assumes that in evaluating actions and attitudes an agent's welfare at all times is equally important. In this chapter, I have two aims. First, I will introduce Parfit's surgery case to show that it is proper for an agent to prefer the situation where her painful experiences are done and over to the situation where she will suffer from pain soon. Second, I will show that whereas the temporal neutrality thesis has a problem in explaining Parfit's case, the care thesis does not have the problem. Given that the care thesis has explanatory power over the temporal neutrality thesis, I will conclude that the former is a more accurate view of prudence than the latter.

This chapter's structure is as follows: In section 3.2, I introduce the surgery case. In particular, I show that whereas the care thesis can explain the intuition in the case, the temporal

neutrality thesis cannot explain the intuition. In section 3.3, I dismiss the contention that the temporal neutrality thesis can also explain the surgery case, so there is no need to endorse the care thesis. In section 3.4 and 3.5, I refute two arguments to show that the surgery case does not threaten the temporal neutrality thesis because the intuition in the case is not reliable. And, in section 3.6, I reply to the claim that the care thesis does not successfully explain the intuition in the surgery case.

3.2 The Surgery Case

While investigating practical rationality, Parfit provides a case where a patient has no memory of her surgery schedule. In particular, Parfit describes this case, which I will name the *surgery case*, as follows:

This kind of surgery is completely safe, and always successful. Since I know this, I have no fears about the effects. The surgery may be brief, or it may instead take a long time. Because I have to cooperate with the surgeon, I cannot have anesthetics. I have had this surgery once before, and I can remember how painful it is. Under a new policy, because the operation is so painful, patients are now afterwards made to forget it. Some drug removes their memories of the last few hours.

I have just woken up. I cannot remember going to sleep. I ask my nurse if it has been decided when my operation is to be, and how long it must take [. . .] She can tell me only that the following is true. I may be the patient who had his operation yesterday. In that case, my operation was the longest ever performed, lasting ten hours. I may instead be the patient who is to have a short operation later today.³¹

³¹ Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), 165.

If the patient went through the operation yesterday, then the patient experienced pain for ten hours. However, given that the patient's painful experiences are done and over in the case where she received the treatment, in terms of prudence it seems proper that the patient prefers the past surgery to the future operation although the surgery in the past is ten times more painful than the surgery in the future.

Theses of prudence must be able to show that, in the surgery case, it is appropriate for the patient to prefer pain in the past to suffering in the future. However, the temporal neutrality thesis has a problem in explaining this intuition in the surgery case. Imagine that agent A is in one of two situations. If agent A is in the first situation, then she receives a painful surgery for ten hours later today. If she is in the second situation, then she has the same surgery for an hour at the same time. In this case, it is prudentially inappropriate that agent A prefers the first situation to the second situation because the former is more painful than the latter. The patient, in the surgery case, is similarly situated to agent A. Just as agent A ought to receive a painful surgery, the patient should go through an operation for one hour or ten hours. The sole difference between agent A and the patient is that the patient's surgeries are set up at different temporal locations. As mentioned above, the temporal neutrality thesis assumes that the timing during which an agent receives advantages/disadvantages is not important in evaluating the agent's actions and attitudes. Therefore, the coherent position for the temporal neutrality thesis is that, just as it is inappropriate for agent A to prefer the ten-hour surgery in the future to the one-hour treatment at the same time, it is also inappropriate for the patient to prefer the ten-hour's past operation to the one-hour's future operation.

Unlike the temporal neutrality thesis, the care thesis can explain why it is appropriate that the patient prefers the past operation to the future surgery. In the surgery case, the past operation

is ten times more painful than the future treatment, so if the patient's welfare at all times is equally important, then it is inappropriate that the patient prefers the past treatment to the future treatment. However, according to the care thesis, if an agent cares for herself at time T1 more than for herself at time T2, then the agent's welfare at T1 is important more than the agent's welfare at T2. I think people care for their future selves more than they care for their past selves. This is why people prefer the situation where they already had painful experiences to the situation where they will suffer from pain. Just as people have this pattern of caring, the patient would also care for herself in the future more than she cares for herself in the past, thus within the framework of the care thesis it is prudentially permissible that the patient cares about her welfare in the future more than she cares about her welfare in the past. In particular, since it is prudentially justifiable that the patient has a partial attitude toward herself in the future, in the surgery case it is prudentially appropriate that the patient prefers the ten-hour's treatment in the past to the one-hour's operation in the future.

It is prudentially appropriate that an agent prefers the scenario where she already had painful experiences to the scenario where she will undergo pain. As shown above, the temporal neutrality thesis has a problem in explaining this phenomenon of pain. Assuming that an agent's welfare at all times is equally important, it is inappropriate that the agent prefers pain in the past to suffering in the past if the past pain is stronger than the future suffering. In contrast, the care thesis can explain why it is appropriate to prefer pain in the past. According to this thesis, an agent's welfare in the future is more important than the agent's welfare in the past. This is why it could be appropriate that an agent prefers the scenario where she experienced strong pain to the scenario where she will have mild suffering. Given that the care thesis has explanatory power over

the temporal neutrality thesis, it is reasonable to say that the former is a good alternative to the latter.

3.3 The Beneficial Attitude Reply

Advocates might contend that there is no need to endorse the care thesis because the temporal neutrality thesis can also explain why in the surgery case it is proper that the patient prefers the past operation. In particular, the temporal neutrality thesis's advocates might utilize the following idea:

‘Since we cannot affect the past, this is a good ground for being less concerned about it.’ [. . .] ‘If we lacked this bias, we would be as much concerned about past pains and pleasures, which we cannot affect. This would distract our attention from future pains and pleasures, which we can affect. Because we would be distracted in this way, we would be less successful in our attempts to get future pleasures and avoid future pains. This would be worse for us.’³²

The idea is that whereas an agent cannot change her past an agent can change her future, so it is prudentially better that an agent cares about future events more than about past events. To put this another way, if an agent has a future bias, then the agent can increase her chance to maximize welfare in her life because the agent would spend more time thinking of her changeable future events than thinking of her unchangeable past events. Based on this idea, advocates of the temporal neutrality thesis might explain the surgery case as follows: The patient has prudential reason to prefer the one-hour surgery in the future to the ten-hour operation in the past because the patient experiences pain just for an hour, and an agent's welfare at all times is equally important. However,

³² Parfit (1984), 168.

it is prudentially permissible that the patient prefers the ten-hour treatment in the past, for the patient has the preference due to her attitude (i.e. the attitude to care about her future events more than about her past events), and this attitude (i.e. a future bias) is conducive to maximizing welfare in her life. Therefore, the temporal neutrality thesis can explain the intuition in the surgery case. In this dissertation, I will name this reply for the temporal neutrality thesis the *beneficial attitude reply*.

I also think it is permissible that an agent cares about her future events more than she cares about her past events. However, I am not sure whether welfare maximization is the reason that allows the attitude to care about future events more than about past events (In particular, I think it is prudentially permissible for an agent to care about future events more than she cares about past events because an agent cares for her future self more than for her past self). Imagine that agent A has stopped pursuing project X because she no longer endorses the project. Agent A has new interests in project Y, and she knows she will keep endorsing this project. In this scenario, assuming that project X is prudentially valuable more than project Y, agent A can maximize her welfare just in case she completes project X. However, if agent A cares about her future events more than about her past events, then it could be the case that agent A completes project Y. For example, agent A might be motivated to pursue project Y because she would consider her future interest (i.e. her interest in project Y) more important than her past interest (i.e. her interest in project X).

The scenario above shows that if an agent cares about future events more than about past events, then it could be the case that the agent pursues her current project rather than completes her previous project, even if the previous project is prudentially valuable more than the current project. Throughout life people encounter a significant number of situations in which they should

decide whether they will keep pursuing previous projects or start pursuing new projects. Therefore, it is not clear to me whether an agent's attitude to care about future events more than about past events is conducive to maximizing welfare in the agent's life. According to the beneficial attitude reply, it is permissible that an agent prefers significant pain in the past to minor suffering in the future because the agent has the preference out of her future bias (i.e. the attitude to care about her future more than about her past), and this bias is conducive to welfare maximization. However, as mentioned above, it is not clear whether this biased attitude brings out a larger amount of advantages for the agent herself, so the beneficial attitude reply does not successfully explain the surgery case.

3.4 The Irrational Regret Reply

To save the temporal neutrality thesis, advocates might take another strategy to dismiss the surgery case. Advocates might say that the intuition in the surgery case is not reliable. In other words, it is problematic that the patient prefers the past surgery, so the fact that the temporal neutrality thesis does not justify the patient's preference does not threaten the thesis. There could be various versions of this strategy depending on why advocates think the preference for the past surgery is inappropriate. Some advocates might utilize Tom Dougherty's argument against future biases. According to Dougherty, future biases are problematic because a future biased agent could regret her previous rational decision. Based on this claim, advocates of the temporal neutrality thesis might claim that agents prefer pain in the past to suffering in the future because they have future biases. Given that future biases are problematic, an agent's preference for past pain is also problematic. Therefore, the fact that the temporal neutrality thesis fails in explaining the surgery

case does not threaten the thesis. In this dissertation, I will name this reply the *irrational regret reply*.

Before examining the irrational regret reply, let me explain why Dougherty thinks future biases are problematic. Dougherty provides the following case to show that future biases are problematic:

Victoria is deciding on Friday how to spend a free hour on her Saturday afternoon. The weather is forecast to be fine, and she is choosing between lazing around in the sun or working in her garden by mowing the lawn. She decides that the value of an orderly lawn is not quite enough to justify forgoing the pleasures of sunbathing for the discomfort of working in the sun [...] But when the hour is up, she notices that the pleasure is now past. And so too would be the painful hour of horticultural exertion. Because she is future-biased, she now discounts the past pleasure of sunbathing and the pain of garden-work. Meanwhile, her evaluation of a pristine lawn remains constant [...] She comes to regret having chosen the sunbathing over the gardening. She thinks, “I wish now that I had chosen differently back then!”³³

Suppose that in terms of prudence Victoria’s decision is rational. The instant benefits from having a sunbath are larger than the long-lasting advantages from having a tidy lawn. According to Dougherty, if Victoria considers what will happen more important than what already occurred, then she could regret having made the rational decision. This is because, due to the previous decision, Victoria has no upcoming advantages (i.e. the advantages that she could have enjoyed if she had mowed her lawn rather than had taken a sunbath). Dougherty contends that this regret is problematic because the previous decision is rational. In particular, Dougherty says that future-biased attitudes are impermissible because if an agent considers what will happen more important

³³ Dougherty, “Future-Bias and Practical Reason” in *Philosophers’ Imprint* 15 (2015), 5.

than what already occurred, then she could regret her decision even in the case where the decision is rational.³⁴

The irrational regret reply assumes that if an agent has a certain attitude (e.g. a future-biased attitude), and this is why the agent regrets her previous rational decision, then the attitude is impermissible. However, for two reasons, I think this assumption for the irrational regret reply is implausible. First, it does not seem to be problematic that an agent regrets her previous decision, even if the previous decision is rational. If Victoria had mowed her lawn, then Victoria would have had upcoming advantages (i.e. the advantages from her mowed lawn). In this situation, given that Victoria prefers the scenario where she has upcoming advantages to the scenario where she has received all benefits, I think there is nothing wrong in Victoria's regret (i.e. the regret that she should have mowed her lawn rather than enjoyed the sunbath). This is because, in Victoria's current perspective, her previous decision's outcome (i.e. the outcome where she has no upcoming benefits) is not satisfactory. To put this another way, if a previous decision's outcome is unsatisfactory in the decision-maker's current perspective, then it seems appropriate that the decision-maker regrets having made the decision even in the case where the previous decision is rational. Unless advocates argue that it is problematic to regret rational decisions, the mere fact that a future-biased agent could regret her rational decision does not seem to threaten future-biased attitudes.

Second, the irrational regret reply's assumption (i.e. an agent's attitude is impermissible if the attitude allows the agent to regret her previous rational decision) is not feasible because, according to this assumption, an agent's moral improvement is problematic. Let me argue first that

³⁴ Joseph Mendola provides a similar objection to future-biases. See Mendola, "Justice Within a Life" in *American Philosophical Quarterly* 41 (2004), 135.

in terms of prudence an agent's moral improvement could be permissible. An agent's moral improvement could bring out advantages for the agent. For instance, the fact that an agent has acquired moral personality can make the agent pleased, or if an agent becomes a morally better person, then the agent can have better relationships (e.g. friendships, parent-child relations, or citizenships) with other agents. However, the main reason why an agent's moral improvement increases the agent's welfare is that an agent's moral improvement is the agent's achievement. An agent makes an achievement if the agent fulfills two conditions.³⁵ First, the agent acquires an intrinsically valuable goal. Second, the agent's efforts are a cause to acquire the intrinsically valuable goal. In particular, as Gwen Bradford argues,³⁶ if an agent fulfills these two conditions, then the fact that the agent satisfies the conditions benefits the agent. This is why in terms of prudence an agent has strong reason to make an achievement even though the processes to make an achievement are painful and burdening. Moral personalities are valuable in themselves. For instance, regardless of whether an agent's altruistic attitude successfully benefits other agents, it is valuable in itself that an agent cares for other agents for their own sake. Therefore, if an agent acquires moral personality as a result of self-training, then the moral improvement benefits the agent.

An agent's moral improvement increases the agent's level of welfare. This is why in terms of prudence an agent's moral improvement could be justifiable. However, the irrational regret reply implies that an agent's moral improvement is problematic because this change in personality can make the agent regret her previous rational decision. Suppose that morally corrupted agent A

³⁵ For more accounts of achievement, see Gwen Bradford, *Achievement* (Oxford: Oxford University Press, 2015); Bradford, "Achievement, Wellbeing, and Value" in *Philosophy Compass* 11 (2016); and Hasko von Kriegstein, "Effort and Achievement" in *Utilitas* 29 (2017).

³⁶ Bradford (2016).

has two options. If agent A engages in discriminative activities, then she has pleasant experiences. If she mows her lawn, then her tidy lawn brings out advantages for her. The discriminative actions bring out the largest amount of benefits for agent A herself, so morally corrupted agent A makes a prudentially rational decision (i.e. the decision to engage in the immoral activities). In this scenario, if agent A acquires the sense of justice, then it could be the case that agent A regrets her prudentially rational decision. Agent A might acknowledge the fact that the discriminative activities brought out the largest advantages for her, but she might wish that she did not engage in the activity because, in her current moral perspective, the immoral pleasures are not valuable at all. Given that agent A could regret her previous rational decision in the case where she acquires moral personality, according to the irrational regret reply, agent A's moral improvement is not permissible even in the case where the improvement is conducive to maximizing welfare in her life.

The irrational regret reply assumes that if an agent has a certain attitude, and this is why the agent regrets her rational decision, then it is problematic that the agent has the attitude. Based on this assumption, advocates of the irrational regret reply contend that future biases are problematic because future-biased agents regret their rational decisions. In particular, since future-biased agents prefer strong pain in the past to mild suffering in the future, preferences for past suffering over future pain are impermissible. This is why Parfit's surgery case does not threaten the temporal neutrality thesis. In this section, I have pointed out that the irrational regret reply's assumption is not intuitively compelling. Furthermore, I have argued that according to this assumption an agent's moral improvement is impermissible although prudence allows an agent to improve herself. Therefore, given that the irrational regret reply's assumption is not plausible, the

irrational regret reply does not successfully save the temporal neutrality thesis from the surgery case.

3.5 The Weak No Regrets Reply

In order to save the temporal neutrality thesis, advocates could utilize another argument against future biases. Preston Greene and Meghan Sullivan argue that future biases are prudentially problematic. In particular, Greene and Sullivan provide the following principle in order to defeat future biases:

Weak No Regrets If an agent has full and accurate information about the effects of the options available to her, then it is rationally permissible for her to avoid options she knows she will regret in favor of ones she knows she will never regret.³⁷

Imagine that agent A has two options. If agent A chooses the first option, then she engages in hedonic activities. If agent A selects the second option, then she does not engage in the activities. Agent A knows that she will have pleasant experiences if she engages in the hedonic activities. However, agent A also knows that she will regret having engaged in the activities because she will live an ascetic life at old age. In this case, which I name the *regret case*, advocates of *Weak No Regrets* would say that, even if agent A loses her chance to maximize welfare, it is permissible that she does not engage in the activities, for she will regret her decision if she engages in the activities.

Greene and Sullivan claim that *Weak No Regrets* is a right understanding of prudence, so if under an assumption this principle has a problematic implication, then it indicates that the

³⁷ Greene and Sullivan, "Against Time Bias" in *Ethics* 125 (2015), 958.

assumption is problematic. Especially, based on this theoretical framework, Greene and Sullivan provide the following case in order to defeat the idea that an agent's future bias is prudentially permissible:

Fine Dining Jack wins a free meal at a fancy French restaurant on Monday morning, and he must schedule the meal for a night sometime in the next week. Given his flexible schedule, every night is equally convenient for him, and there are no other considerations that would make the meal more enjoyable or more likely to occur one night rather than another.³⁸

In this case, Jack receives the same meal regardless of when he visits the restaurant. Based on this fact, Greene and Sullivan claim that Jack's decision to postpone his meal schedule is irrational because he has no reason to delay the schedule. Furthermore, Greene and Sullivan contend that if Jack is future-biased, then *Weak No Regrets* justifies Jack's decision to have the meal on the last day of the event. Therefore, they say, future biases are problematic because if an agent has this attitude, then a correct understanding of prudence (i.e. *Weak No Regrets*) has a problematic implication.

Advocates might claim that, since the argument above successfully shows the problem of future bias, Parfit's surgery case fails in defeating the temporal neutrality thesis. In this dissertation, I will name this reply for the temporal neutrality thesis the *Weak No Regrets Reply*. Before examining the weak no regrets reply, let me explain why Greene and Sullivan think that the principle of *Weak No Regrets* justifies Jack's decision to postpone his meal schedule. According to Greene and Sullivan, if future-biased Jack has the meal earlier than the last day of free meal event, then Jack regrets his decision. This is because he prefers the scenario where he will obtain

³⁸ *Ibid.*, 959.

benefits to the scenario where he already received advantages, and he has no upcoming pleasant experience as a result of having the meal. In contrast, if Jack has the meal on the last day, then he does not regret the decision because he has no chance to enjoy the meal after the day. Given that *Weak No Regrets* allows an agent to choose the option which she will not regret, this principle justifies Jack's decision to postpone his schedule even though he receives the same meal at any of the days.

For discussion, I will assume that in the case where Jack has a future-biased attitude, *Weak No Regrets* justifies Jack's decision to postpone his schedule. Even under this assumption, however, the weak no regrets reply does not succeed in supporting the temporal neutrality thesis. For the weak no regrets reply saves the thesis just in case Greene-Sullivan's argument shows the problem of future biases, but their argument fails in defeating future biases. Greene and Sullivan claim that Jack's decision to postpone his meal schedule is irrational because Jack has no reason to postpone the schedule. This claim is difficult to accept. Unlike Greene-Sullivan's claim, the fact that Jack has no reason to postpone the schedule does imply that Jack's decision to delay the schedule is irrational. Imagine that Jack has two foods in front of him. The food on his left side is as good as the food on his right side, and he can have just one of them. In this case, Jack's decision to have any of the foods is rational although he has no reason to choose one over the other. Especially, the fact that the foods on the dishes are equally makes Jack's decision to choose any of the dishes rational. If one says that Jack's decision is irrational because he has no reason to prefer one over the other, then this claim is not plausible. In the same vein, even if Jack does not have reason to delay his meal schedule, it does not follow that Jack's decision to postpone the schedule is irrational.

Another problem of Greene-Sullivan's argument is that, in *Fine Dining*, Jack has reason to postpone his meal schedule. Greene-Sullivan's argument relies on the premise that Jack has no reason to postpone his meal schedule because he receives the same quality of food regardless of when he visits the restaurant. However, this premise is not a correct understanding of *Fine Dining*. Regret is a frustrated desire accompanied by negative feelings. In other words, the fact that an agent regrets her decision means that the agent wishes she did not make the decision, and the agent experiences negative feelings in regard to the decision. According to the desire-fulfillment theory, the fact that an agent's desire is frustrated decreases the agent's level of welfare. Furthermore, in any plausible theory of welfare, an agent's negative feelings decrease the agent's welfare. Therefore, if an agent regrets her previous decision, then the agent's level of welfare decreases. In particular, given that prudence requires an agent to increase the agent's own welfare, an agent has prudential reason not to choose an option if she will regret the choice. According to Greene and Sullivan, if Jack has the free meal earlier than the last day, then Jack regrets his decision because as a result of the decision he has no more chance to enjoy the free meal. In this scenario, due to the fact that he will regret his decision, Jack has prudential reason to postpone his meal schedule. This is why Greene-Sullivan's premise (i.e. Jack has no reason to delay the free meal schedule) is incorrect.

The last problem of Greene-Sullivan's argument is that *Weak No Regrets* has the implication which defeats the temporal neutrality thesis. According to *Weak No Regrets*, if an agent's future self takes a negative attitude (i.e. regret) to choosing option X, but does not take a negative attitude to selecting option Y, then it is prudentially permissible that the agent chooses option Y even in the case where option X brings out larger advantages for the agent herself. To put this another way, an agent's attitude in the future has a decisive role in evaluating the agent's

decision. However, as I will argue below, an agent's attitude in the past does not have the importance which *Weak No Regrets* assigns to an agent's attitude in the future. In particular, even if an agent's past self takes a negative attitude to option X, it is impermissible that the agent chooses option Y if option X brings out larger advantages for the agent. This asymmetry between past and future attitudes could endorse the idea that a past self's prudential status is lower than a future self's prudential status. Since a past self's status is lower than a future self's status, a past self's negative attitude is less important than a future self's negative attitude. Assuming that a past self has a lower status to a future self, it is plausible to say that an agent's past welfare is less important than the agent's future welfare. Therefore, proponents cannot utilize Greene-Sullivan's argument because *Weak No Regrets* has the implication which threatens the temporal neutrality thesis.

Let me show the point above with an example. Agent B knew that she would lose her ascetic aspiration. Furthermore, agent B also knew that after she loses the aspiration, hedonic activities would bring her a significant amount of advantages. However, agent B wished that she would not engage in hedonic activities because in her current perspective hedonistic activities were inappropriate. As agent B expected, agent B loses her interests in ascetism. Furthermore, she has a chance to engage in hedonic activities. In this scenario (hereafter I will name this case the *wish case*), it could be the case that agent B has a reason not to engage in hedonic activities because she did not want herself to engage in the activities. However, this reason (i.e. agent B did not want herself to engage in hedonic activities) does not seem to have a decisive role in determining what agent B should do. In particular, given that hedonic activities increase agent B's level of welfare, and agent B does not endorse ascetism anymore, in terms of prudence it is rational that agent B engages in hedonic activities. For instance, if agent B refuses to engage in hedonic activities just

because she once disliked the activities, then agent B's decision is irrational in the perspective of prudence.³⁹

In the wish case, agent B's past self takes a negative attitude to the situation where agent B engages in hedonic activities. However, agent B does not have strong reason to forego hedonic activities. In contrast, according to *Weak No Regrets*, an agent's negative attitude in the future has decisive role in evaluating the agent's decision in the present. Especially, even if hedonic activities bring out the largest advantages for an agent, it is prudentially permissible that the agent does not engage in the activities if the agent's future self takes a negative attitude to the scenario where the agent engages in the activities. Therefore, if *Weak No Regrets* is assumed, then it follows that an agent's negative attitude in the past has minor importance, but an agent's negative attitude in the future has significant importance. As mentioned above, this implication can support the idea that an agent's past self has lower status than the agent's future self. In particular, this idea can justify the claim that past welfare is less important than future welfare. As this discussion shows, Greene-Sullivan's argument threatens the temporal neutrality thesis, so advocates cannot utilize this argument in order to save their understanding of prudence (i.e. the temporal neutrality thesis) from Parfit's case.

3.6 The Motivation Objection

³⁹ Krister Bykvist and Elizabeth Harman also endorse the position that if an agent either no longer endorses ideals or does not endorse ideals yet, then it is prudentially irrational that the agent sacrifices extra advantages for the ideals. See Bykvist, "Prudence for Changing Selves" in *Utilitas* 18 (2006); and Harman, "'I'll be Glad I Did It' Reasoning and the Significance of Future Desires" in *Philosophical Perspectives* 23 (2009).

In the surgery case, it is appropriate that the patient prefers the scenario where she already received the ten-hour's surgery to the scenario where she will have the one-hour's operation. As argued above, the temporal neutrality thesis has a problem accommodating this phenomenon, and this is why I think the care thesis is a more accurate view of prudence than the temporal neutrality thesis. However, critics could claim that the care thesis also does not successfully explain the surgery case. In particular, based on this claim, critics could conclude that though the temporal neutrality thesis is a limited understanding of prudence, the care thesis is not a good alternative to the temporal neutrality thesis. Let me explain why some critics could say that the care thesis does not successfully explain the surgery case. Imagine that agent A has made significant influences on agent B's mental states. In this scenario (hereafter I will call this scenario the *region case*), if agent A cares for agent B because she has made influences on agent B's mental states, then it is justifiable that agent A considers agent B's welfare more important than a stranger's welfare. However, if agent A is motivated to care just because agent B came from the region where she used to live, then agent A's partial attitude is inappropriate even though agent A is psychologically related to agent B.

The scenario above shows that, depending on why an agent cares for other agents, the agent's partial attitudes toward the other agents could be impermissible. In particular, critics of the care thesis might contend that if an agent's psychological relation (or friendship) is what makes the agent's caring important, then the agent should be motivated to care due to her psychological relation (or friendship). In other words, if an agent's psychological relation is the normative ground for her caring, then the agent should be motivated to care due to the normative ground. If the agent is motivated by other considerations, then the fact that the agent cares for other agents does not justify the agent's partial attitudes toward the other agents. This is why, in the region case, if agent

A cares for agent B just because agent B came from the region where she used to live, then agent A's partial attitude toward agent B is impermissible. The normative ground for agent A's caring is agent A's psychological relation with agent B. Therefore, agent A's caring justifies agent A's partial attitude just in case agent A cares for agent B because she is psychologically related to agent B.

In the surgery case, the patient cares for herself in the future more than she cares for herself in the past because the former is located in the future, and the latter is located in the past. If the patient was motivated to care, for instance, due to her psychological relations with herself in the past and future, then the patient would have cared for herself at the times to the same degree. This is because the patient's mental states are equally related to the patient's mental states in the past and future. Critics of the care thesis might contend that, in the surgery case, the patient's psychological relation is what makes the patient's caring important. However, the patient cares for herself in the future due to her future self's temporal location. Therefore, the patient's caring does not justify the patient's partial attitude toward her future self. In particular, since the patient's caring does not justify the partial attitude toward herself in the future, the care thesis cannot explain the intuition in the surgery case. Hereafter I will name this objection the *motivation objection*. I also think an agent's motivation to care is important. This is why, in the region case, if agent A cares for agent B just because agent B came from the region where she used to live, then agent A's caring does not justify agent A's partial attitude toward agent B. However, the motivation objection does not successfully defeat the care thesis because, besides normative grounds for caring (e.g. psychological relation or friendship), other facts could be appropriate motivations to care.

Two interpretations are possible in regard to the region case. The first interpretation is that an agent's caring has significance just in case the agent cares for other agents due to her caring's normative grounds (e.g. the fact the agent is psychologically related to the other agents, or the fact that the agent is friends with the other agents). In the region case, if agent A cares for agent B just because agent B came from the region where agent A used to live, then agent A's caring is not significant because agent A's motivation to care is not that she is psychologically related to agent B. Agent A's caring is significant if, and only if, agent A cares for agent B due to her caring's normative ground (i.e. her psychological relation). The second interpretation is that the region of origin is not a proper reason to care. If agent A cares for agent B just because agent B came from a certain region, then agent A's caring is not important because in deciding what to do it is not important where agent B came from. According to this second interpretation, unless agent A cares for agent B due to the fact that agent B came from a certain region, agent A's caring could be significant.

Critics of the care thesis endorse the first interpretation in order to refute the care thesis. However, there is no reason to accept the first interpretation over the second interpretation. In fact, the second interpretation is better than the first interpretation because the former has explanatory power over the latter. Suppose that, in *My Division*, agent X cares for agent Y and Z because agent X believes that agent Y and Z are her, not someone else. In this case, it is permissible that agent X cares about agent Y and Z's welfare more than about other people's welfare. The second interpretation (i.e. unless an agent cares for other agents just because the other agents came from the region where she used to live, the fact that the agent cares for the other agents could justify the agent's decision for the other agents) can explain this phenomenon. According to this interpretation, the fact that agent X cares for agent Y and Z justifies agent X's partial attitude

toward agent Y and Z because agent X's motivation to care is not that agent Y and Z came from certain regions. However, the first interpretation (i.e. unless an agent cares for other agents due to her caring's normative grounds, the agent's caring is insignificant) has a problem to explain why it is justifiable that agent X cares about agent Y and Z's welfare more than she cares about other people's welfare. Agent X's motivation to care is not her caring's normative ground (i.e. the fact that she is psychologically related to agent Y and Z), so the first interpretation implies that agent X's caring is insignificant. In other words, within the framework of the first interpretation, the fact that agent X cares for agent Y and Z does not justify agent A's partial attitude toward agent Y and Z.

The discussion above shows that, even if an agent's motivation to care is not the caring's normative grounds (e.g. friendship or psychological relation), it does not follow that the agent's caring does not have significance. Based on this fact of caring, let me argue that the motivation objection does not successfully defeat the care thesis. In the surgery case, the patient cares for herself in the future more than for herself in the past because the former is located in the future, and the latter is located in the past. However, the fact that the patient cares for her future self due to the self's temporal location does not imply that the patient's caring is not significant. This is because, as the discussion above shows, an agent's caring can have significance even in the case where the agent's motivation to care is not the caring's normative grounds. Unless agent A cares for agent B just because agent B came from a certain region, agent A's caring can have significance. In particular, the fact that it is appropriate for the patient to prefer the past surgery indicates that, although spatial location is not a proper motivation to care, temporal location is an appropriate motivation to care. Therefore, the care thesis successfully explains Parfit's surgery case. Moreover,

due to this explanatory strength, the care thesis is a good alternative to the temporal neutrality thesis.

3.7 Conclusion

In this chapter, I had two goals. First, I argued that the care thesis can explain why it is appropriate for an agent to prefer pain in the past to suffering in the future, but the temporal neutrality thesis cannot explain this phenomenon. Second, I dismissed three attempts to save the temporal neutrality thesis, and I replied to a criticism of the care thesis. Given that the care thesis has explanatory power over a widely endorsed thesis (i.e. the temporal neutrality thesis), and main attempts to save the temporal neutrality thesis do not successfully support the thesis, it is reasonable to say that the care thesis is a plausible understanding of the self-regarding domain: Prudence. In other words, the care thesis is a good alternative to the dominant view of prudence (i.e. the temporal neutrality thesis).

CHAPTER 4

NEAR-TERM BIAS

4.1 Introduction

In chapter 2, I suggested a new thesis of prudence: The care thesis. According to the care thesis, if an agent cares for herself at time T1 more than for herself at time T2, then it is permissible that the agent cares about her welfare at T1 more than about her welfare at T2. In chapter 3, I argued that the care thesis has explanatory power over a widely endorsed thesis of prudence (i.e. the temporal neutrality thesis). The temporal neutrality thesis has a problem in explaining why it is appropriate to prefer pain in the past to suffering in the future. According to the temporal neutrality thesis, it is inappropriate that an agent prefers strong pain in the past to mild suffering in the future because an agent's welfare at all times is equally important. On the other hand, the care thesis can explain why in terms of prudence it is justifiable to prefer pain in the past. Within the framework of the care thesis, it is justifiable that an agent prefers strong pain in the past because if an agent already went through painful experiences, then her cared-for intrapersonal friend (i.e. her future self) will not experience pain. In this last chapter, I will address a concern about the care thesis. Critics could say that the care thesis is problematic because the thesis justifies an irrational attitude: A near-term

bias. I will argue that near-term biases are rational, so the above concern does not refute the care thesis.

The structure of this chapter is as follows: First, I explain the near-term bias objection to the care thesis. In particular, I point out that the near-term bias objection successfully defeats the care thesis just in case near-term biases are prudentially impermissible. Second, I introduce five arguments against near-term biases. Especially, I show that these arguments fail in showing that near-term biases are prudentially impermissible. Based on the fact that main arguments fail in refuting near-term biases, I conclude that the near-term bias objection does not threaten the care thesis.

4.2 The Near-Term Bias Objection

Critics might say that the care thesis justifies an agent's near-term bias (i.e. an agent's attitude to care about what happens in the near future more than about what occurs in the far future). For instance, if an agent cares for her near-future self more than for her far-future self, then within the framework of the care thesis it is prudentially permissible that the agent has a near-term bias. This is because a near-term biased agent would end up benefiting her strongly cared-for self (i.e. herself in the near future) rather than her weakly cared-for self (i.e. herself in the far future). Critics might contend that an agent's near-term bias is prudentially impermissible. For instance, if an agent has a near-term bias, then the agent might enjoy minor benefits right now even in the case where she can receive significant advantages a few minutes later. Therefore, critics might conclude that the care thesis is a wrong view of prudence because this thesis justifies a prudentially problematic

attitude (i.e. a near-term bias). In this dissertation, I will name this objection the *near-term bias objection*.

In the literature of prudence, it is widely assumed that near-term biases are irrational.⁴⁰ For instance, in the voice of Socrates, Plato says an agent should choose larger benefits regardless of whether the agent obtains the advantages in the near future or far future. In particular, Plato writes as follows:

[Y]ou put the pleasures together and the pains together, both the near and the remote, on the balance scale, and then say which of the two is more [...] If you weigh the pleasant things against the painful, and the painful is exceeded by the pleasant – whether the near by the remote or the remote by the near – you have to perform that action in which the pleasant prevails [...] While the power of appearance often makes us wander all over the place in confusion, often changing our minds about the same things and regretting our actions and choices with respect to things large and small, the art of measurement in contrast, would make appearances lose their power by showing us the truth, would give us peace of mind firmly rooted in the truth, and would save our life.⁴¹

Given that near-term biases are considered irrational, at a glance the care thesis seems to be problematic because this thesis justifies a prudentially irrational attitude. However, the mere fact that near-term biases are considered irrational is not strong enough to refute the care thesis. Advocates of the care thesis can reply that widely endorsed views could be wrong. In particular, assuming that an agent cares for herself in the near future more than for herself in the far future, it

⁴⁰ For instance, see Plato, *Protagoras*, trans. Stanley Lombardo and Karen Bell (Indianapolis: Hackett Publishing Company, 1992), 356b-e; Adam Smith, *The Theory of Moral Sentiments* (Oxford: Oxford University Press, 1976), VI.i.11; Henry Sidgwick, *The Methods of Ethics* (London: Macmillan, 7th ed., 1907), 381; Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 32-33; Meghan Sullivan, *Time Biases: A Theory of Rational Planning and Personal Persistence* (Oxford: Oxford University Press, 2018); and Dale Dorsey, “A Near-Term Bias Reconsidered” in *Philosophy and Phenomenological Research* 99 (2019).

⁴¹ Plato (1992), 356b-e.

is reasonable to say that an agent's near-term bias is rational because a near-term biased agent would end up benefitting her strongly cared-for self (i.e. herself in the near future). Therefore, in order to defeat the care thesis, critics should show that an agent's near-term bias is prudentially irrational.

4.3 The Arbitrariness Argument

Meghan Sullivan suggests an argument to show that near-term biases are impermissible. In particular, Sullivan thinks that an agent's near-term bias is impermissible due to the following reason:

- (1) At any given time, a prudentially rational agent's preferences are insensitive to arbitrary differences. (The Non-Arbitrariness principle)
- (2) Relative distance from the present is an arbitrary difference between events.
- (3) If you are near-biased, your preferences are sensitive to when an event is scheduled relative to the present.

C. So at any given time, near-biased preferences are not rationally permissible.⁴²

In this argument, Sullivan makes three assumptions. First, in prudential deliberation, an agent should not consider prudentially irrelevant factors. Second, if an agent has a near-term bias, then the agent considers whether an event occurs in the near future or in the far future. Third, it is prudentially irrelevant whether events occur in the near future or in the far future. Therefore, Sullivan concludes that in terms of prudence an agent's near-term bias is impermissible because

⁴² Sullivan (2018), 37-38.

this bias makes an agent consider prudentially irrelevant factors (i.e. an event's relative temporal distance from the present). I will name this argument against near-term biases the *arbitrariness argument*.

The arbitrariness argument relies on the assumption that, in evaluating events, it is not important whether the events occur in the near future or in the far future. However, advocates of the care thesis have a reasonable ground to refuse this assumption. According to the care thesis, if agent A cares for herself at time T1 more than she cares for herself at time T2, then it is permissible that agent A cares about her welfare at T1 more than about her welfare at T2. This is because, due to the fact that agent A cares for herself at T1 more than for herself at T2, agent A's friendship with the T1-self is valuable more than her friendship with the T2-self. Furthermore, since agent A is psychologically related to herself at T1, the fact that agent A strongly cares for herself at T1 justifies her partial attitude toward her T1-self. In the same vein, if an agent cares for herself in the near future more than for herself in the far future, then in terms of prudence it is permissible that the agent cares about her welfare in the near future more than she cares about her welfare in the far future.

Assuming that it is permissible to care about welfare in the near future more than about welfare in the far future, it is important whether events occur in the near future or far future. For instance, if an agent can bring out beneficial events, and the events increase the agent's welfare in the near future, then it could be permissible that the agent brings out the events, even if she loses her chances to benefit herself in the far future. Similarly, if an agent undergoes an operation either in the near future or in the far future, then in terms of prudence it could be appropriate that the agent prefers the scenario where she receives the treatment in the far future to the scenario where she goes through the treatment in the near future. Therefore, if an agent cares for herself in the

near-future more than for herself in the far future, then the agent's near-term bias is permissible because due to this bias the agent is sensitive to prudentially relevant facts (i.e. the facts of whether events occur in the near future or in the far future). In other words, the arbitrariness argument does not successfully show that an agent's near-term bias is impermissible because if an agent cares for herself in the near future more than for her far-future, then the relative distance from the present is important.

4.4 The Instrumental Irrationality Argument

Besides the arbitrariness argument, Sullivan suggests another argument to show that an agent's near-term bias is a problematic attitude. In particular, Sullivan provides her second argument as follows:

(1) At any given time, a rational agent prefers that her life going forward go as well as possible. (The Success principle)

(2) If you are near-biased, then in distant tradeoffs you will prefer and choose the present, lesser good over the greater, future good.

(3) Your life going forward would go better if you preferred and chose the greater, future good in a distant tradeoff.

C. So near-biased preferences are not rationally permissible, insofar as you face a distant tradeoff choice.⁴³

⁴³ *Ibid.*, 23-24.

If an agent wants to have a cup of hot tea, but orders a cup of ice tea, then the agent's decision is irrational. This is because the decision is not conducive to obtaining what the agent wants (i.e. having a cup of hot tea). According to the argument above (hereafter I will name this argument the *instrumental irrationality argument*), an agent's near-term bias is irrational because an agent's near-term bias is not conducive to obtaining what the agent wants: Her life going well as much as possible.

Just as the instrumental irrationality argument assumes, I also think a near-term biased agent might lose her chance to have the prudentially best future. To put this another way, since a near-term biased agent might choose minor advantages in the near future rather than significant benefits in the far future, a near-term biased agent might not be able to maximize welfare in her future. Therefore, assuming that an agent wants to have the prudentially best future, at a glance the instrumental irrationality argument seems to show that near-term biases are irrational. However, for two reasons, I think the instrumental irrationality argument fails in defeating near-term biases. First, an agent's wish to have the best future could be conditional. In particular, if an agent has a near-term bias, then the agent's wish for the best future is conditional, and this conditional wish is compatible with the agent's near-term bias. Second, the fact that an agent's near-term bias is not conducive to living the best future does not necessarily imply that near-term biases are irrational. This fact could imply that the wish for the best future is irrational, so an agent should revise the wish.

Let me show the first limit of the instrumental irrationality argument. People have other-regarding wishes. In particular, many people want that everyone in the world lives the prudentially best life unless the person obtains advantages from immoral activities. Pointing out these people who wish for everyone's best life, one could contend that these people's family-oriented biases are

irrational in the sense that family-oriented biases are not conducive to maximizing welfare in the world. I think this contention against family-oriented biases is plausible only in certain cases. To put this another way, even if people have this other-regarding wish (i.e. the wish that everyone lives the best life), it could be the case that people's family-oriented biases are not irrational in the sense above. This is because what people actually want could be that, unless their family members receive disadvantages, everyone obtains the largest amount of advantages. In particular, the fact that people prefer minor benefits for their families to significant advantages for strangers indicates that their wishes to maximize welfare in the world are conditional. In the same vein, an agent's wish to maximize future welfare could be conditional in that the agent wants to maximize future welfare just in case her near-future self does not lose significant advantages. In particular, provided that an agent considers her near-future events more important than her far-future events, it is reasonable to say that the agent's fully elaborated wish is that she obtains the largest benefits in the future unless she loses significant benefits in the near future. Therefore, even if a near-term biased agent chooses benefits in the near future rather than larger advantage in the far future, it does not defeat near-term biases. This is because what a near-term biased agent actually wants is that she maximizes welfare in the future unless she loses a significant amount of advantages in the near future.

Even if an agent's wish to have the best future is unconditional (i.e. an agent unconditionally wants to obtain the largest amount of benefits in the future), it does not still follow that the agent's near-term bias is irrational. Let me show the second limit of the instrumental irrationality argument with the following analogy: Imagine that agent A wants everyone in the world to receive the largest advantages. However, agent A also wants to give some gifts to her parents, go on trips with her partner, and offer good qualities of education to her children. In

particular, agent A has this family-oriented bias because she cares for her family members more than for strangers. In this case, one could claim that agent A's family-oriented bias is irrational because, due to this bias, agent A loses her chances to obtain what she wants: Maximizing welfare in the world. However, unlike this objection to agent A's family-oriented bias, the fact that the bias is not conducive to welfare maximization does not necessarily imply that the bias is irrational. Given that agent A cares for her family members more than for strangers, it is also reasonable to contend that agent A's other-regarding wish (i.e. the wish that everyone lives the best life) is problematic. In particular, rather than eliminating the family-oriented bias, agent A should revise her other-regarding wish in order to make the wish compatible with agent A's desire to take care of her family.

The fact that an agent's other-regarding wish (i.e. the wish that everyone in the world lives the prudentially best life) is not compatible with the agent's family-oriented bias does not show that the agent's family-oriented bias is irrational. In the same vein, the fact that an agent's near-term bias is not compatible with the agent's wish to have the best future does not show that the agent's near-term bias is irrational. Imagine that agent A wants it to be the case that her life goes as well as possible. However, agent A also prefers the scenario where she obtains minor benefits in the near future to the scenario where she receives significant advantages in the far future. Especially, agent A has this bias because she cares for her near-future self more than for her far-future self. In this case, advocates of the instrumental irrationality argument might contend that agent A's near-term bias is irrational because due to this temporal bias agent A fails in obtaining what she wants: Obtaining the largest amount of benefits in the future. However, advocates of near-term biases could reply that agent A's wish to obtain the largest benefits is irrational. This is because if agent A makes decisions just to bring out the largest benefits in her future, then she

might fail in benefitting her cared-for self (i.e. herself in the near future). This reply for near-term biases is as plausible as the instrumental irrationality argument, so critics cannot defeat near-term biases just referring to the fact that an agent's near-term bias is incompatible with the agent's wish to have the best future. In other words, unless critics show that the instrumental irrationality argument is plausible more than the above reply for near-term biases, critics cannot utilize the argument.

4.5 The Foreseen Regret Argument

Besides Sullivan's arguments, there are two more arguments against near-term biases. These arguments utilize the concept of regret to defeat the justifiability of near-term biases. The first regret-based argument is introduced by Derek Parfit. Parfit describes his regret-based argument as follows:

You do not *now* regret your bias towards the near. But you *will*. When you pay the price – when you suffer the pain that you postponed at the cost of making it worse – you will wish that you did not care more about your nearer future. You will regret that you have this bias. It is irrational to do what you know you will regret.⁴⁴

The argument above (hereafter I will name this argument the *foreseen regret argument*) makes two assumptions. First, it is irrational that an agent performs an action if the agent knows she will regret having performed the action. Second, agents know they will regret having had near-term biases. Based on these claims, the foreseen regret argument concludes that near-term biases are irrational.

⁴⁴ Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), 187.

One could claim that the foreseen regret argument does not successfully defeat near-term biases because this argument's second assumption (i.e. agents know they will regret having had near-term biases) could be false. For example, one might say it is possible that an agent does not regret having had a near-term bias because the agent has a tendency not to regret her previous decisions. Furthermore, it is possible that an agent does not know whether she will regret having had a near-term bias because the agent does not have clear understandings of her near-term biased decision's outcomes. I also think there are cases where the foreseen regret argument's second assumption is false, so this argument does not show that every near-term bias is problematic. In other words, even within the framework of this argument, some near-term biases could be permissible. However, the fact that the foreseen regret argument does not refute every near-term bias is not conducive to saving the care thesis. Critics might contend that if an agent knows she will regret having had a near-term bias, then the agent's near-term bias is irrational. However, even in this scenario, the care thesis justifies an agent's near-term bias if the agent cares for her near-future self more than she cares for her far-future self. This is why the care thesis is a problematic view.

The care thesis implies that if an agent cares for herself in the near future more than for herself in the far future, then the agent's near-term bias is rational even in the case where the agent knows she will regret having had the temporal bias. However, this implication does not refute the care thesis because the foreseen regret argument's first assumption is not true. To put this another way, even if an agent knows she will regret having performed an action, it does not follow that it is irrational for the agent to perform the action. Imagine that agent A, who is a hedonist, has two options. If agent A chooses the first option, then agent A engages in hedonic activities throughout life, and these activities bring out tremendous advantages for agent A. If agent A chooses the

second option, then she loses her chance to engage in the activities, so she cannot live the best life for her. While deliberating what to do, a fortuneteller lets agent A know that at the end of life she will become ascetic, and she will regret having chosen the first option (i.e. the option to engage in hedonic activities). In this scenario, agent A has a prudential reason not to engage in the hedonic activities because the regret will decrease her last moment's welfare. However, given that agent A receives tremendous benefits, it seems rational for agent A to choose the first option. In other words, agent A's reason to engage in the activities (i.e. the activities benefit her) seems to be stronger than her reason not to participate in them (i.e. the regret will decrease her future welfare). Therefore, it is rational that agent A engages in hedonic activities in order to enjoy larger advantages.

According to the foreseen regret argument, an agent's near-term biased decision is irrational in the case where the agent knows she will regret the decision. In contrast, even if an agent knows that she will regret her near-term bias, the care thesis implies that the agent's near-term biased decision is rational. This is why, according to the foreseen regret argument, the care thesis is a problematic view of prudence. In the case above, agent A knows that if she engages in hedonic activities, then she will regret her decision. However, agent A's decision to engage in hedonic activities is rational. In particular, agent A should engage in the hedonic activities because due to the activities she can live the prudentially best life. This case shows that though the care thesis justifies an agent's near-term bias even in the case where the agent knows she will regret having had the temporal bias, this fact does not cause a problem to the thesis. Therefore, the foreseen regret argument does not successfully show that the care thesis is a problematic view of prudence.

4.6 The Awry Agency Argument

Dale Dorsey also suggests a regret-based argument against near-term biases. Dorsey contends as follows:

To take a concrete case, imagine that at time t I am offered one delicious lollypop at t_1 or two delicious lollypops at t_2 . At t_2 , I regret my t -self's near-term biased decision to have one lollypop at t_1 . But if, while maintaining this attitude of regret for my t -decision, I make the *very same decision* to have one lollypop at t_3 rather than two at t_4 , it would seem that something has gone horribly awry in my prudential agency [...] The problem with the near-term bias is that this *form* of prudential rationality *allows* this combination of attitudes in those who display the typical form of past-directed negative reactive attitudes.⁴⁵

To put this another way, according to Dorsey, if an agent has a near-term bias, then it could be the case that the agent makes a near-term biased decision while she regrets having made a similar decision before. This is why, Dorsey contends, near-term biases are irrational in regard to prudence. Hereinafter I am going to name Dorsey's argument against near-term biases the *awry agency argument*.

Before examining the awry agency argument, let me explain what makes it irrational that an agent has a near-term bias while she regrets having had a near-term bias. Dorsey contends as follows:

In displaying past-directed reactive attitude, one seems to be taking on a form of intrapersonal address: address to one's past self. And the content of such an address is very similar to the content of similar forms of interpersonal address. In interpersonal address of this kind, one demands that

⁴⁵ Dorsey, "A Near-Term Bias Reconsidered" in *Philosophy and Phenomenological Research* 99 (2019), 471.

another person take one's own perspective seriously in their deliberation. In intrapersonal address the upshot is similar: one is recognizing the failure to live up to an intrapersonal demand that one's past self take one's current perspective seriously [...] But when combined with NTB [near-term bias] the person taking such an attitude *both* holds that one's past self should have taken one's current perspective more seriously, while *refusing* to take seriously a similar demand on behalf of one's (far) future self – a refusal which is a direct result of the near-term bias.⁴⁶

Suppose that agent A makes a near-term biased decision while she regrets having made a similar decision before. According to Dorsey, this state is problematic because, in this state, agent A refuses to fulfill her future self's demand (i.e. agent A should consider her far future as important as her near future) while she makes a similar demand to her past self (i.e. the past self should have considered the past self's far future as important as the self's near future). Dorsey calls this combination of attitudes *unsavory*. In particular, Dorsey says near-term biases are not permissible because an agent's near-term bias could allow the agent to have the combination of unsavory attitudes.

I endorse the awry agency argument in that if an agent regrets her previous near-term bias, then the agent's current near-term bias is not rational. In particular, as Dorsey points out, this state is problematic because, in the state, an agent refuses to fulfill a demand (i.e. the agent should consider her far future as important as her near future) while she makes a similar demand to her past self (i.e. the past self should have considered the self's far future as important as the self's near future). However, I disagree with the idea that the fact that a near-term biased agent can have the unsavory combination of attitudes makes every near-term bias problematic. If an agent regrets her previous near-term bias, then her current near-term bias is irrational, but if an agent does not

⁴⁶ *Ibid.*

regret her previous near-term bias, then the agent's current near-term bias could be prudentially rational. In particular, since the care thesis can accommodate this fact of near-term biases, the awry agency argument does not successfully show that the care thesis is a wrong understanding of prudence.

Let me provide an analogy to show that if an agent does not regret her previous near-term bias, then the agent's current near-term bias could be permissible. Imagine that agent A benefits her friend rather than a stranger (i.e. agent B) even though she can bring out larger benefits for the stranger. After agent A completes her beneficial acts, agent B blames agent A saying that agent A should not have considered who possible beneficiaries are. In particular, agent B says people should choose their beneficiaries just considering who receives larger benefits. While blaming agent A, agent B has a chance to benefit either her friend or a stranger. Although agent B can bring out larger advantages for the stranger, agent B benefits her friend because agent B prioritizes her friend over a stranger. In this case, agent B shows two attitudes. Agent B demands that agent A consider a stranger's benefits as important as a friend's advantages, but she refuses to consider a stranger's benefits as important as her friend's advantages. The combination of these two attitudes (i.e. making certain demands to other agents, but refusing to live up the agent's own life following the demands) is unsavory. This is why, I think, agent B's decision for her friend is not permissible. Furthermore, given that agent B has unsavory attitudes because of her partial attitude toward her friend, it is reasonable to contend that agent B's partial attitude toward her friend is also impermissible.

The case above raises a question of whether the fact that agent B's partial attitude is problematic implies that every partial attitude toward a friend is inappropriate. A possible answer, which I will name the *pro-impartiality account*, is that since a friendship-biased agent can show

unsavory attitudes, every partial attitude toward a friend is problematic. In contrast, according to another possible answer (hereafter I will name it the *pro-partiality account*), the fact that agent B's attitude is problematic just implies that an agent's partial attitude toward a friend is inappropriate if the agent blames other agents for having partial attitudes toward their friends. I think the pro-partiality account is more accurate than the pro-impartiality account because, in many cases, morality allows partial attitudes toward friends. Imagine that agent C endorses people's partial attitudes toward their friends. Especially, agent C thinks that in terms of morality it is appropriate to prioritize a friend's benefits over a stranger's advantages because this partial attitude is a proper response to the valuable relationship. While having this thought, agent C has a chance to benefit either her friend or a stranger. Agent C cares for her friend more than she cares for a stranger, so agent C makes a decision for her friend even though she can bring out larger benefits for the stranger.

In the case above, agent C's decision is permissible unless the stranger desperately needs agent C's help. To put this another way, given that agent C does not blame other agents for having partial attitudes toward their friends, a theory of morality must be able to justify agent C's partial attitude toward her friend. The pro-partiality account can explain why agent C's partial attitude is appropriate. According to this account, since agent C does not blame agent A, agent C's partial attitude could be justifiable. On the other hand, the pro-impartiality account implies that agent C's attitude is problematic, for an agent can have the unsavory combination of attitudes if the agent has a partial attitude toward the agent's friend. Given that the pro-partiality account has explanatory power over the pro-impartiality account, it is reasonable to say that the former is a more accurate understanding of friendship than the latter. To put this another way, advocates of

the pro-impartiality account should accept that the pro-partiality account is at least as plausible as their position.

As shown above, if agent B blames agent A for being partial to agent A's friend, then it is problematic that agent B considers her friend's benefits more important than a stranger's advantages. However, the fact that a friendship-biased agent could have an unsavory combination of attitudes does not imply that every partial attitude toward a friend is problematic. In the same vein, it is reasonable to claim that if an agent reproaches her past self for having had a near-term bias, then it is problematic that the agent has a near-term bias, but the fact that a near-term biased agent could have the unsavory combination of attitudes does not make every near-term bias problematic. This claim about near-term biases can be re-stated with the concept of regret. If an agent regrets having had a near-term bias, then it is problematic that the agent considers her near future more important than her far future. However, the fact that a near-term biased agent could have the unsavory combination of attitudes (i.e. regretting her near-term bias but still having the same temporal bias) does not make every near-term bias problematic. In particular, if an agent does not regret her previous near-term bias, then her current near-term biased decision could be permissible.

Advocates of the care thesis can accommodate the above phenomenon of near-term biases (i.e. the phenomenon that if an agent does not regret her previous near-term bias, then it could be permissible that the agent has a near-term bias, but if an agent regrets her previous near-term bias, then the agent's near-term bias is inappropriate). Imagine that agent A cares for herself in the near future more than she cares for herself in the far future. In this case, if agent A does not blame her past self for the past self's near-term bias, then agent A's near-term bias is permissible. This is because, due to agent A's friendship and psychological relation with the near-future self, agent

A's caring justifies agent A's near-term bias. However, just as an agent's friendship-bias is impermissible in the case where the agent reproaches others for their friendship-biases, agent A's near-term bias is not permissible if agent A blames her past self for the self's near-term bias. Given that the care thesis makes right verdicts regarding the rationality of near-term biases, it is reasonable to say that the awry agency argument does not successfully show the care thesis's problem.

4.7 The Extreme Bias Argument

The last argument targets a certain form of near-term bias. Critics might contend that if an agent has an extreme form of near-term bias, then the agent might sacrifice entire far-future advantages just to enjoy minor near-future advantages. For instance, if an agent strongly cares about what happens in the near future, but she is totally indifferent to what occurs in the far future, then the agent might decide to enjoy a sandwich in the near future at the cost of tremendous advantages in the far future. Critics might contend that, even if there are some cases where near-term biases are prudentially justifiable, in any circumstances this extreme temporal bias should not be justified. However, according to the care thesis, an agent's extreme form of near-term bias could be permissible if an agent strongly cares for herself in the near future, but the agent is totally indifferent to herself in the far future. This is because an agent's extreme form of near-term bias will bring out advantages for the agent's strongly cared-for self (i.e. the agent herself in the near future). Therefore, critics might conclude that the care thesis is an incorrect understanding of prudence. In this dissertation, I will name this argument to refute the care thesis the *extreme bias argument*.

Just as the extreme bias argument assumes, I also think that an agent's extreme form of near-term bias is impermissible. However, this argument does not successfully defeat the care thesis because the care thesis would predict that, regardless of how much an agent cares for herself in the near future, if an agent sacrifices tremendous far-future benefits just to enjoy minor near-future advantages, then the decision is irrational. In particular, advocates of the care thesis would say that an agent's decision to sacrifice tremendous far-future benefits is irrational, for an agent's reason to choose tremendous far-future benefits is stronger than the agent's reason to select minor near-future advantages. Imagine that agent A can benefit herself in the far future more than she can benefit herself in the near future. In this case, other things being equal, agent A's reason to choose the option for her far-future self (i.e. this option benefits the far-future self) is stronger than agent A's reason to select the option for her near-future self (i.e. this option benefits the near-future self). For agent A can benefit her far-future self more than she benefits her near-future self. However, according to the care thesis, if agent A cares for herself in the near future more than for herself in the far future, then this pattern of caring enhances agent A's reason to choose the option for her near-future self. In particular, depending on scenarios, agent A's enhanced reason to choose the option for her near-future self could be stronger than agent A's reason to select the option for her far-future self. For instance, if agent A cares for her near-future self much more than for her far-future self, and agent A can bring out minor extra benefits for her far-future self, then agent A's enhanced reason for her near-future self could be stronger than agent A's reason for her far-future self.

Though caring enhances the strengths of reasons, caring cannot make a reason to choose minor benefits (i.e. the fact that the choice increases the agent's own welfare) stronger than a reason to select enormous advantages (i.e. the fact that the selection benefits the agent herself). In

the domain of morality, even if an agent cares for her friend much more than for a stranger, the agent is required to benefit the stranger if she can bring out minor benefits (e.g. instant pleasure from a sandwich) for her friend but enormous advantages (e.g. good health) for the stranger. In the same vein, regardless of how much an agent cares for her near-future self, if an agent can bring out minor benefits for her near-future self but enormous advantages for her far-future, then the agent's reason to choose the enormous advantage would be stronger than the agent's reason to select the minor benefits. Therefore, even within the framework of the care thesis, it is irrational that an agent sacrifices enormous far-future benefits to enjoy minor near-future benefits, regardless of whether she cares for her near-future self much more than for her far-future self. In the same vein, the care thesis implies that an agent's extreme form of near-term bias is impermissible because if an agent has this bias, then the agent could sacrifice enormous benefits for minor advantages.

Besides the fact that an agent's reason to choose enormous far-future benefits is stronger than the agent's reason to select minor near-future advantages, the care thesis has another ground to predict that an agent should sacrifice minor benefits in the near future. If an agent cares for herself in the near future, then the care thesis would predict that the agent should sacrifice near-future advantages for larger far-future benefits because as a result of the sacrifice the agent could live a prudentially better life in the near future. As mentioned in chapter 1, if an agent sacrifices her advantages at time T1, and the sacrifice brings out advantages for the agent herself at time T2, then the sacrifice is meaningful in that the sacrifice at T1 is conducive to obtaining advantages at T2. In particular, the fact that the sacrifice is meaningful contributes to making the agent at T1 live a better life for herself. This idea of sacrifice does not necessarily imply that all things considered an agent's meaningful sacrifice makes the agent live a better life at the moment of sacrifice. In

order to decide whether an agent's sacrifice makes the agent herself live a better life, the benefits from a meaningful sacrifice should be weighed with the sacrificed advantages. If the former is larger than the latter, then all things considered the agent lives a better life at the moment of sacrifice.

Imagine that agent A has two options. If agent A chooses the first option, then she can enjoy a sandwich in the near future but loses a tremendous amount of advantages in the far future. In contrast, if agent A selects the second option, then she sacrifices her chance to have a sandwich in the near future but can enjoy tremendous advantages in the far future. In this case, agent A has two sources to benefit herself in the near future. The first source is the sandwich. When agent A has the sandwich, she would feel instant pleasure from the sandwich. The second source is the meaningful sacrifice. If agent A sacrifices the sandwich to enjoy tremendous advantages in the far future, then the fact that the sacrifice is meaningful contributes to making agent A live a prudentially better life in the near future. The benefits from a sandwich are not huge because the sandwich gives agent A just instant pleasure. On the other hand, the meaningful sacrifice has significant prudential value because as a result of the sacrifice agent A can enjoy great advantages. This is why agent A can live the prudentially best life in the near future when she sacrifices her sandwich. As this example shows, if an agent sacrifices minor benefits in the near future to bring out enormous benefits in the far future, then not only does the agent maximize welfare in the far future but the agent also let herself in the near future live the best life. Therefore, if an agent cares for her near-future self, then the agent should sacrifice near-future benefits to cause huge far-future advantages.

The discussion above shows that, even if an agent cares for her near-future self much more than for her far-future self, for two reasons the care thesis would not allow the agent to sacrifice

enormous far-future advantages just to enjoy minor near-future advantages. First, even if the fact that an agent cares for her near-future self enhances the agent's reason to select minor near-future benefits, this reason is overwhelmed by the agent's reason to choose tremendous far-future advantages. Second, an agent can live the best life in the near future when she sacrifices minor near-future benefits for tremendous far-future advantages. Given that the care thesis does not allow an agent to sacrifice enormous far-future benefits, this thesis would not justify an agent's extreme form of near-term bias. This is because if an agent has this bias, then the agent might choose minor near-future benefits rather than huge far-future advantages. Therefore, it is reasonable to say that the extreme bias argument does not successfully defeat the care thesis, for the extreme bias argument relies on the assumption that the care thesis justifies an agent's extreme form of near-term bias.

4.8 Conclusion

In the literature of prudence, it is widely believed that near-term biases are impermissible. Based on this belief of prudence, critics suggest the near-term bias objection to the care thesis. According to the near-term bias objection, the care thesis is a wrong view of prudence because this thesis justifies a prudentially impermissible attitude: A near-term bias. However, the mere fact that the care thesis justifies near-term biases does not threaten the care thesis. This is because a widely endorsed belief could be false, and the care thesis has a plausible ground to justify an agent's near-term bias. In order to defeat the care thesis, advocates of the near-term bias objection must be able to explain why near-term biases are not permissible. In this chapter, I introduced five arguments against near-term biases. Especially, I argued that these arguments do not successfully defeat the temporal bias. Therefore, given that main arguments do not refute near-term biases, I think it is

reasonable to say that the near-term bias objection fails in show that the care thesis is a wrong view of prudence.

CHAPTER 5

CONCLUSION

In this dissertation, I investigated whether in the domain of prudence an agent's welfare at all times is equally important. The temporal neutrality thesis assumes that an agent's welfare at every moment of life has the same importance in evaluating actions and attitudes. In particular, according to the personal identity argument, an agent's welfare in the past, present, and future has the same significance because an agent is herself regardless of when she exists. I argued that this argument does not successfully endorse the temporal neutrality thesis. It is true that personal identity is a ground for an agent to benefit herself at other temporal locations. However, personal identity is not the sole relation which is prudentially significant. An agent has intrapersonal friendships with herself at other times. Furthermore, an agent is psychologically related to herself in the past and future. I argued that these intrapersonal relations have prudential significance, and due to these relations, it could be permissible for an agent to care about her welfare at certain times more than she cares about her welfare at other times. In particular, I suggested the care thesis of prudence, according to which if agent A cares for herself at time T1 more than she cares for herself at time T2, then agent A's welfare at T1 is important more than agent A's welfare at T2. This is because agent A's pattern of caring makes agent A's friendship with her T1-self more valuable than agent A's friendship with her T2-self, and due to agent A's psychological relation with her T1-self, the

fact that agent A cares for her T1-self is prudentially significant in evaluating agent A's actions and attitudes.

The care thesis has explanatory power over the temporal neutrality thesis. As Parfit's surgery case shows, it is prudentially appropriate for an agent to prefer the scenario where she already experienced strong pain to the scenario where she will have mild suffering. The temporal neutrality thesis encounters a problem in explaining Parfit's surgery case. Assuming that the timing during which an agent experiences pain is irrelevant in evaluating the agent's preference, it follows that an agent should prefer weak pain in the future to strong suffering in the past. In contrast, the care thesis can explain Parfit's surgery case. According to the care thesis, it is appropriate that an agent prefers strong pain in the past to minor suffering in the future, for an agent cares for herself in the future more than for herself in the past. Critics might contend that, even though the care thesis has a strength in that the thesis can explain the above phenomenon of pain, the care thesis is a wrong view of prudence because this view justifies a prudentially impermissible attitude: A near-term bias. I dismissed this objection pointing out that main arguments against near-term biases do not successfully refute near-term biases. Given that the care thesis has explanatory power over a widely endorsed thesis (i.e. the temporal neutrality thesis), and a main objection fails in defeating the care thesis, it is reasonable to conclude that the care thesis is a feasible understanding of prudence.

BIBLIOGRAPHY

- Archard, David. "The Obligations and Responsibilities of Parenthood" in *Procreation and Parenthood: The Ethics of Bearing and Rearing Children*, eds. David Archard and David Benatar (Oxford: Oxford University Press, 2010).
- Aristotle. *Nicomachean Ethics*, ed. Roger Crisp (Cambridge: Cambridge University Press, 2005).
- Bradford, Gwen. *Achievement* (Oxford: Oxford University Press, 2015).
- Bradford, Gwen. "Achievement, Wellbeing, and Value" in *Philosophy Compass* 11 (2016).
- Bradford, Gwen. "Problems for Perfectionism" in *Utilitas* 29 (2017).
- Bramble, Ben. "A New Defense of Hedonism about Well-Being" in *Ergo: An Open Access Journal of Philosophy* 3 (2016).
- Brink, David. "Prospects for Temporal Neutrality" in *The Oxford Handbook of Philosophy of Time*, ed. Craig Callender (Oxford: Oxford University Press, 2011).
- Bruckner, Donald W. "Present Desire Satisfaction and Past Well-Being" in *Australasian Journal of Philosophy* 91 (2013).
- Bykvist, Krister. "Prudence for Changing Selves" in *Utilitas* 18 (2006).
- Bykvist, Krister. "Comments on Dennis McKerlie's 'Rational Choice, Changes in Values over Time, and Well-Being'" in *Utilitas* 19 (2007).
- Dancy, Jonathan. *Ethics without Principles* (Oxford: Oxford University Press, 2004).
- Dorsey, Dale. "The Significance of a Life's Shape" in *Ethics* 125 (2015).
- Dorsey, Dale. "A Near-Term Bias Reconsidered" in *Philosophy and Phenomenological Research* 99 (2019).
- Dougherty, Tom. "Future-Bias and Practical Reason" in *Philosopher's Imprint* 15 (2015).
- Feldman, Fred. *Pleasure and the Good Life: Concerning the Nature, Varieties and Plausibility of Hedonism* (Oxford: Clarendon Press, 2004).
- Fletcher, Guy. "Taking Prudence Seriously" in *Oxford Studies in Metaethics: Volume 14*, ed. Russ Shafer-Landau (Oxford: Oxford University Press, 2019).
- Fletcher, Guy. "A Fresh Start for the Objective-List Theory of Well-Being" in *Utilitas* 25 (2013).
- Glasgow, Joshua. "The Shape of a Life and the Value of Loss and Gain" in *Philosophical Studies* 162 (2013).

- Greene, Preston and Sullivan, Meghan. “Against Time Bias” in *Ethics* 125 (2015).
- Harman, Elizabeth. ““I’ll be Glad I Did It” Reasoning and the Significance of Future Desires” in *Philosophical Perspectives* 23 (2009).
- Heathwood, Chris. “Desire-Fulfillment Theory” in *The Routledge Handbook of the Philosophy of Well-Being* (London: Routledge, 2016).
- Heathwood, Chris. “Subjective Theories of Well-Being” in *The Cambridge Companion to Utilitarianism*, eds. Ben Eggleston and Dale E. Miller (New York: Cambridge University Press, 2014).
- Helm, Bennett. “Friendship” in *The Stanford Encyclopedia of Philosophy*, Fall Edition (2017), URL = <https://plato.stanford.edu/archives/fall2017/entries/friendship/>.
- Hurka, Thomas. *Perfectionism* (Oxford: Oxford University Press, 1993).
- Hurka, Thomas. “The Justification of National Partiality” in *Drawing Morals: Essays in Ethical Theories* (New York: Oxford University Press, 2011)
- Jaworska, Agnieszka. “Caring and Internality” in *Philosophy and Phenomenological Research* 74 (2007).
- Jeske, Diana. “Friendship, Virtue, and Impartiality” in *Philosophy and Phenomenological Research* 57 (1997).
- Kauppinen, Antti. “Meaningfulness and Time” in *Philosophy and Philosophical Research* 84 (2012).
- Keller, Simon. “Four Theories of Filial Duty” in *The Philosophical Quarterly* 56 (2006).
- Keller, Simon. *Partiality* (Princeton: Princeton University Press, 2013).
- Kolodny, Niko. “Love as Valuing a Relationship” in *Philosophical Reviews* 112 (2003).
- Lazar, Seth. “The Justification of Associative Duties” in *Journal of Moral Philosophy* 13 (2016).
- Mendola, Joseph. “Justice Within a Life” in *American Philosophical Quarterly* 41 (2004).
- Murphy, Mark C. “The Simple Desire-Fulfillment Theory” in *Noûs* 33 (1999).
- Nagel, Thomas. *The Possibility of Altruism* (Princeton: Princeton University Press, 1970).
- Nozick, Robert. *Anarchy, State, and Utopia* (New York: Basic Books, 1974).
- Parfit, Derek. *Reasons and Persons* (Oxford: Clarendon Press, 1984).
- Plato. *Protagoras*, trans. Stanley Lombardo and Karen Bell (Indianapolis: Hackett Publishing Company, 1992).
- Purves, Duncan. “Desire Satisfaction, Death, and Time” in *Canadian Journal of Philosophy* 47 (2017).
- Scheffler, Samuel. “Relationships and Responsibilities” in *Philosophy and Public Affairs* 26 (1997).

- Seidman, Jeffrey. "Valuing and Caring" in *Theoria* 75 (2009).
- Sidgwick, Henry. *The Methods of Ethics* (London: Macmillan, 7th ed., 1907).
- Slot, Michael. "Goods and Times" in *Goods and Virtues* (Oxford: Oxford University Press; 1983).
- Smith, Adam. *The Theory of Moral Sentiments* (Oxford: Oxford University Press, 1976).
- Sullivan, Meghan. *Time Biases: A Theory of Rational Planning and Personal Persistence* (Oxford: Oxford University Press, 2018).
- Telfer, Elizabeth. "Friendship" in *Proceeding of the Aristotelian Society* 71 (1971).
- Temkin, Larry. *Rethinking the Good* (Oxford: Oxford University Press, 2012).
- Trebilcot, Joyce. "Aprudentialism" in *American Philosophical Quarterly* 11 (1974).
- Velleman, David. "Well-Being and Time" in *Pacific Philosophical Quarterly* 72 (1991).
- Von Kriegstein, Hasko. "Effort and Achievement" in *Utilitas* 29 (2017).