

# Does the Ability to Do Otherwise Include a Disposition?

## A Defense of the New Dispositionalism

Yoshinari HATTORI

### 0. Introduction

This presentation presents a defense of new dispositionalism (ND) in the contemporary theory of moral responsibility, which was developed by Vihvelin (2004, 2013) and Fara (2008). According to ND, the ability required for moral responsibility can be understood as a disposition, which is a metaphysical concept. ND has the advantage of dealing well with two significant problems with the so-called ability to do otherwise. First, ND enables a compatibilist concept of the ability to do otherwise against van Inwagen's consequence argument. Second, ND answers Frankfurt's challenge, a thought experiment intended to show that the ability to do otherwise is not required for moral responsibility. Despite these advantages, the application of dispositions alone is insufficient for the ability to do otherwise. An additional pair of conditions, one epistemic and the other metaphysical, should be added to address further problems.

### 1. Basic Idea of ND

Let us begin with a story. Imagine an agent who has done something morally wrong, was coerced to do it by another's threats or became obsessed with the idea of doing it due to mental illness. In this setting, we would not say that this agent is morally responsible. Why? We might expect the answer that doing the immoral act was the only option available to her then. She was unable to do otherwise. The concept of ability referred to in this answer is usually called the ability to do otherwise or alternative possibility, and its requirement for moral responsibility is referred to as the principle of alternative possibility (PAP). As in the example given here, PAP has an intuitive pull. However, to validate this principle, the concept of the ability to do otherwise should be analysed. Vihvelin and Fara, the two ND advocates, propose an analysis of this concept as a disposition.<sup>1</sup> The core idea is as follows:

#### Standard ND Analysis

Suppose that agent *S* does not do *X* in world *w*. Agent *S* has the ability to do *X* in *w* iff *S* has the (intrinsic) disposition to do *X* when *S* tries to do *X*. (Vihvelin, 2004, 431, 438; Fara, 2008, 848)

---

<sup>1</sup> In this paper, we do not discuss the ability to perform mental action. For those interested in this kind of ability, see Smith (2003), who obtains the analyses of the ability to believe or desire otherwise by disposition.

Here, the term ‘disposition’ is used in a metaphysical sense. It refers to the way that an object behaves or changes in response to the existence of a certain situation. Here, the object’s behavior or change is called *manifestation*, and the situation in which it exhibits manifestation is called *stimulus*. A disposition is individuated by a pair of a manifestation and a stimulus referred to in the form ‘disposition to M when C’. Fragility is taken as a paradigm case of a disposition, understood as the disposition to break when dropped or shocked (Choi, 2006; Choi & Fara, 2021; Lewis, 1997; Manley & Wasserman, 2011).

Dispositions have three other characteristics. First, a dispositional property is usually conceived of as *intrinsic*. An intrinsic property is such that, if  $x$  has it, then any of  $x$ ’s perfect physical duplicates  $x^*$  also has it (Marshall & Weatherson, 2018). The instantiation of an intrinsic property should not depend on the existence of something external to its bearer. Although some philosophers deny that all dispositions are intrinsic, this paper exclusively focuses on intrinsic dispositions.<sup>2</sup>

Second, a dispositional property need not always exhibit its manifestation when its stimulus obtains. Two types of things can stop a disposition from exhibiting its manifestation. First, a *fink* is something that removes a disposition from its bearer when a stimulus obtains (Lewis, 1997, 147–148; Martin, 1994, 2–4). Suppose that a sorceress observes a fragile glass, and when it is about to be dropped, she changes its molecular structure so that it is no longer fragile and does not break. Here, she functions as a fink. Second, a *mask* is something that prevents the manifestation of a disposition when a stimulus obtains but does not remove the disposition itself (Bird, 1998, 228; Johnston, 1992, 233). Suppose that a fragile glass is packed with some material, and consequently it does not break easily, even when dropped. Here, the packing material is a mask.

Third, if a disposition is to be ascribed to someone or something, it is necessary that a wide range of actual or hypothetical circumstances should be taken into account. Even if a glass does not break in a certain circumstance of being dropped because it is packed with some material, this does not mean that it is not fragile. If a glass breaks in a wide range of possible dropping circumstances—that is, if it is not packed with the material in most of them—it is possible to ascribe fragility to the glass.

## 2. Two Advantages of ND

ND’s proposal regarding the ability to do otherwise has two main advantages. First, it can reject van Inwagen’s consequence argument and secure the possibility of a compatibilist concept of the ability to do otherwise. Second, it can uphold PAP, which is intuitively attractive, against Frankfurt’s challenge. To understand these advantages, let us start with an overview of basic concepts in the theory of moral responsibility.

Two main groups of positions are found in accounts of moral responsibility: *compatibilism* and *incompatibilism*, defined by the concept of *determinism*. Suppose that  $P(t, w)$  expresses a proposition that represents the complete state of a possible world  $w$  at moment  $t$  and that  $L(w)$  expresses the conjunction of all the laws of physics in  $w$ . Then, determinism is defined as follows:

### Determinism

---

<sup>2</sup> Mckittrick (2003) opposes this dominant view. Park (2017) further objects to Mckittrick.

Determinism is true in  $w$  iff for any moment  $t$  and  $t'$  (after  $t$ ), the conjunction of  $P(t, w)$  and  $L(w)$  entails  $P(t', w)$ .

This form of determinism was first set forth by Peter van Inwagen (1975, 186), and it is still the one most frequently referred to by contemporary theorists, including ND advocates (Fara, 2008, 860; Fisher & Ravizza, 1998, 14; Lewis, 1981, 113; McKenna & Coates, 2021; Smith, 2003; Vihvelin, 2013, 1).

With reference to the above definition, compatibilism claims that moral responsibility is compatible with the truth of determinism, and incompatibilism claims that it is not.

One major issue between these two camps is the question whether or not the ability to do otherwise is compatible with the truth of determinism. Van Inwagen (1983, 93–105) develops an argument that supports the incompatibility of alternative possibility with determinism. He terms this the consequence argument. Although he presents several versions, in the following, I here reconstruct what he calls the third argument (van Inwagen, 1983, 93–105).<sup>3</sup>

Let  $N$  be a sentential operator such that  $N\alpha$  reads ‘ $\alpha$  and no one is, or ever has been, able to make it the case that not- $\alpha$ .’<sup>4</sup> In the logic of the  $N$ -operator, the following two inference rules are valid.

- ( $\alpha$ )  $\Box p \vdash Np$
- ( $\beta$ )  $Np, N(p \supset q) \vdash Nq$

Suppose that time  $t$  is the moment before anyone was born in some world  $w$  and that  $t'$  is the moment at which someone acts in a certain way in  $w$ . The following two premises seem plausible.

- (1)  $NP(t, w)$
- (2)  $NL(w)$

Furthermore, suppose that determinism is true in  $w$ . Then, by the definition of determinism:

- (3)  $\Box(P(t, w) \& L(w) \supset P(t', w))$

We can infer the following:

- (4)  $\Box(P(t, w) \supset (L(w) \supset P(t', w)))$  (by (3) and the Exportation Rule)
- (5)  $N(P(t, w) \supset (L(w) \supset P(t', w)))$  (by (4) and ( $\alpha$ ))
- (6)  $N(L(w) \supset P(t', w))$  (by (1), (5) and ( $\beta$ ))
- (7)  $NP(t', w)$  (by (2), (6) and ( $\beta$ ))

<sup>3</sup> The consequence argument was first proposed in (van Inwagen, 1975, 191). While the third argument is valid, the first version is not unless the inference rules about the  $N$ -operator are admitted.

<sup>4</sup> This reading of the  $N$ -operator is adopted from Fara (2008, 862).

Conclusion (7) reads ' $P(t', w)$  and no one is, or ever has been, able to make it the case that not- $P(t', w)$ .' This means that any agent does not have the ability to perform actions other than the ones they do in fact perform. Therefore, the consequence argument suggests that the truth of determinism deprives agents of the ability to do otherwise.

If physicists someday discover the truth of determinism in the actual world, the consequence argument entails that people do not have the ability to do otherwise than they do. If PAP is also true, it follows that people are not morally responsible for their actions.

ND is not concerned about this possibility. Its advocates argue that, so long as the ability to do otherwise is analysed by a disposition, the rule ( $\beta$ ) in the consequence argument can be rejected, and this ability is compatible with determinism.<sup>5</sup> Admittedly, even under ND's interpretation of ability, (1)  $NP(t, w)$ , (2)  $NL(w)$ , and (5)  $N(P(t, w) \supset (L(w) \supset P(t', w)))$  are true because agents do not have the disposition to change the past, the laws of nature, or the definition of determinism. However, (7) is underivable under the interpretation of ND. Suppose that in a deterministic world, Susan does not raise her hand. Even so, it is perfectly possible for her to have the disposition to raise her hand if she tried. Her brain is awake and is working as usual, and the nerves that connect her arm to the overall nervous system are not cut off—these facts suffice for her possession of the appropriate disposition and ability to raise her hand. Following ND's proposal, then, rule ( $\beta$ )  $Np, N(p \supset q) \vdash Nq$  is to be rejected, and the ability to do otherwise is compatible with determinism.

The second advantage of ND is that it responds to the challenge posed by the Frankfurt-style case, a thought experiment intended to show that PAP is false (Frankfurt, 1969, 835–836). A Frankfurt-style case would run something like the following: consider an era in which neuroscientific technology has developed far in advance of what could be foreseen today. In such an era, Elizabeth plans to shoot the mayor of New York City. Black, a neurosurgeon, who longs for the death of the mayor, is aware of this plan. He wants Elizabeth to accomplish it on her own, but he is concerned that she might waver in pulling the trigger. Hence, Black stealthily implants a device into her brain as Elizabeth is sleeping the night before the planned murder. This device will detect any neuronal signal indicating hesitation to pull the trigger, and it is programmed to manipulate her to kill nevertheless if it detects this signal. The next day, Elizabeth kills the mayor as she planned. The device did not detect any hesitation signal.

In this story, the device was a counterfactual intervener: it would have intervened in the production of Elizabeth's action if it had detected a certain neuronal signal, but it did not do so. Because the device did not intervene, Elizabeth's action in pulling the trigger was under her control. Most people would agree that she is morally responsible for her action. Meanwhile, this device may undermine the intuition that she had the ability not to pull the trigger because, even if she had not decided to pull the trigger, she could not have avoided doing so, due to the device. Frankfurt concludes that PAP is false.

ND advocates deny that the Frankfurt-style case deprives an agent of the ability to do otherwise. According to them, the device is a fink or mask. Since a dispositional

---

<sup>5</sup> For another objection to the rule ( $\beta$ ), see McKay & Johnson (1996).

property is an intrinsic property as we saw, the device does not affect whether Elizabeth has some disposition. Hence, it is possible that Elizabeth has the disposition, hence the ability, to stop pulling the trigger when she tries to do so, but the device would remove this disposition or prevent it from becoming manifested were she to try to do so.

We have seen two advantages of ND. First, the ability to do otherwise, according to ND, is irrelevant to determinism. Therefore, a physical discovery as to whether the actual world is deterministic would not have any impact on ND advocates' concerns regarding whether people have such an ability. Second, the ability to do otherwise, according to ND, is secured even in the Frankfurt-style case. Therefore, PAP is not falsified, under this interpretation.

### 3. A New ND Analysis of the Ability to Do Otherwise

It may be accepted that the ability to do otherwise includes a disposition. However, the further question arises: do such abilities consist only of disposition? To address this question, the ND analysis of the ability to do otherwise should be modified, as follows:

#### New ND Analysis

Suppose that agent *S* does not do *X* in world *w*. *S* has the ability to do *X* in *w* iff  
 (Dispositional Condition) *S* has the (intrinsic) disposition to do *X* when *S* tries to do *X*;  
 (Epistemic Condition) if what *S* does in *w* is the omission of *X* and not another action instead of *X*, *S*'s not doing *X* in *w* is not motivated by *S*'s belief that there are finks or masks of the disposition; and  
 (Metaphysical Condition) no finks or masks operate in *w* to stop the disposition from becoming manifested.

In the rest of this paper, I make the case why the epistemic and metaphysical conditions given above should be added to the analysis of the ability to do otherwise.

Whittle (2010) objects to the standard ND analysis that the dispositional concept of the ability to do otherwise is not the type of ability that Frankfurt and his followers are challenging. Since the ascription of a disposition requires consideration of a wide range of possible circumstances, if *S* has the disposition to do otherwise, it follows that she is able to do across a wide range of circumstances. This concept of ability is called *global* ability. Whittle claims that it is the *local* ability to do *X*—that is, the ability to do *X* in a particular circumstance—that is challenged in the Frankfurt-style case (Whittle, 2010, 2). In the given Frankfurt-style case, Elizabeth has the global ability to stop pulling the trigger because she is able to do so across a wide range of possible circumstances—the device is not implanted into her brain in most of them; however, she does not have the ability to do perform a different act in the particular circumstance where the device is implanted. Whittle claims that the ability to do otherwise should be interpreted as a local ability. To establish this, she considers a case where an agent does not have the ability to do otherwise than it does but nevertheless has the disposition to do so. Suppose that Ben, an excellent swimmer, has been forcibly bound to a chair with a rope. He helplessly watches as a child drowns in a lake. In this case, we cannot say that Ben is morally responsible for not helping the child. However, Whittle attacks the standard ND analysis, claiming that Ben has the disposition to help the child if he tries

to do so, though the rope functions as a mask for this disposition. She concludes that the disposition, or global ability, is irrelevant to the ability to do otherwise.

This objection shows, at most, that the dispositional condition alone is not sufficient for the ability to do otherwise. The epistemic and metaphysical conditions in my proposal can clarify why Ben does not have the ability to help the child, depending on the interpretation of the given story.

First, as Ben helplessly watches the child drown in the lake, he might not try to help the child. In a case such as this, the epistemic condition explains why Ben does not have the ability to help the child. The fact that Ben does not try to help the child may be motivated by his belief that being tied by a rope makes it impossible for him to do so. This epistemic condition implies that he is not morally responsible.

Let me comment on the if-clause in the epistemic condition. In Ben's case, in world  $w$ , he does not do anything at all instead of helping the child; he simply omits to help the child, as he is restrained by the rope. Imagine that Elizabeth in the Frankfurt-style case knows about the implantation of the device, and she pulls the trigger because she believes that there is no way for her not to do so, due to the device. In this case, Elizabeth performs the action of pulling the trigger instead of not pulling it. Is Elizabeth morally responsible for pulling the trigger? I have no clear answer or any intuition in response to this question. Therefore, it is safe to restrict the epistemic condition to cases where agents omit to do action  $X$  and not to extend it to cases where agents perform some action instead of doing  $X$ .

Second, as Ben helplessly watches the child drown in the lake, he could try to help the child, for example, by trying to tear through the rope. In this case, the metaphysical condition explains why he does not have the ability to help the child. This is because the rope in  $w$  has the function of a mask to prevent his disposition from being manifested.

#### 4. Conclusion

This paper defends ND, the idea that the ability to do otherwise includes the disposition to do otherwise as well. This has two advantages: (1) securing PAP against the Frankfurt-style case and (2) rejecting the consequence argument of van Inwagen.

However, the standard ND analysis is still imperfect because it cannot deal with Ben's case, where an agent does not have the ability to do otherwise but does have the disposition to do so. The origin of this problem is the claim in the standard ND analysis that the ability to do otherwise consists only of a disposition. However, this paper proposes that epistemic and metaphysical conditions should be added to the dispositional condition. These conditions offer a solution to the second problem.

#### References

- Bird, A. 1998. "Dispositions and Antidotes." *Philosophical Quarterly* 48, no. 191: 227–34. <https://doi.org/10.1111/1467-9213.00098>.
- Choi, S. 2006. "The Simple vs. Reformed Conditional Analysis of Dispositions." *Synthese* 148, no. 2: 369–79. <https://doi.org/10.1007/s11229-004-6229-z>.
- Choi, S., and M. Fara. 2021. "Dispositions." In *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta (Edition). <https://plato.stanford.edu/archives/spr2021/entries/dispositions/>.

- Fara, M. 2008. "Masked Abilities and Compatibilism." *Mind* 117, no. 468: 843–65. <https://doi.org/10.1093/mind/fzn078>.
- Fisher, J. M., and M. Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Frankfurt, H. G. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66, no. 23: 829–39. <https://doi.org/10.2307/2023833>.
- Johnston, M. 1992. "How to Speak of the Colors." *Philosophical Studies* 68, no. 3: 221–63. <https://doi.org/10.1007/BF00694847>.
- Lewis, D. 1981. "Are We Free to Break the Laws?." *Theoria* 47, no. 3: 113–21. <https://doi.org/10.1111/j.1755-2567.1981.tb00473.x>.
- Lewis, D. 1997. "Finkish Dispositions." *Philosophical Quarterly* 47, no. 187: 143–58. <https://doi.org/10.1111/1467-9213.00052>.
- Manley, D., and R. Wasserman. 2011. "Dispositions, Conditionals, and Counterexamples." *Mind* 120, no. 480: 1191–227. <https://doi.org/10.1093/mind/fzr078>.
- Marshall, D., and B. Weatherson. 2018. "Intrinsic vs. Extrinsic Properties." In *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta (Edition). <https://plato.stanford.edu/archives/spr2018/entries/intrinsic-extrinsic/>.
- Martin, C. B. 1994. "Dispositions and Conditionals." *Philosophical Quarterly* 44, no. 174: 1–8. <https://doi.org/10.2307/2220143>.
- McKay, T. J., and D. Johnson. 1996. "A Reconsideration of an Argument Against Compatibilism." *Philosophical Topics* 24, no. 2: 113–22. <https://doi.org/10.5840/philtopics199624219>.
- McKenna, M., and D. J. Coates. 2021. "Compatibilism." In *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta (Edition). <https://plato.stanford.edu/archives/fall2021/entries/compatibilism/>.
- McKittrick, J. 2003. "A Case for Extrinsic Dispositions." *Australasian Journal of Philosophy* 81, no. 2: 155–74. <https://doi.org/10.1080/713659629>.
- Smith, M. 2003. "Rational Capacities, or: How to Distinguish Reclessness, Weakness, and Compulsion." In *Weakness of Will and Practical Irrationality*, edited by S. Stroud, and C. Tappolet: 17–38. Oxford University Press.
- Park, S. 2017. "Against Extrinsic Dispositions." *Review of Contemporary Philosophy* 16: 92–103. <https://doi.org/10.22381/RCP1620174>.
- van Inwagen, P. 1975. "The Incompatibility of Free Will and Determinism." *Philosophical Studies* 27, no. 3: 185–99. <https://doi.org/10.1007/BF01624156>.
- van Inwagen, P. 1983. *An Essay on Free Will*. Clarendon Press.
- Vihvelin, K. 2004. "Free Will Demystified: A Dispositional Account." *Philosophical Topics* 32, no. 1: 427–50. <https://doi.org/10.5840/philtopics2004321/211>.
- Vihvelin, K. 2013. *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. Oxford University Press.
- Whittle, A. 2010. "Dispositional Abilities." *Philosopher's Imprint* 10, no. 12: 1–23.