

## THE DIF\*

Our topic is the difference variously labeled as that between, on the one hand, “doing” or “causing” or “making” something happen and, on the other, “allowing” or “letting” it happen. The reality and importance of this difference—which we will refer to as *the Dif*—are recognized in law and morality, it being nearly universally acknowledged that causing harm is *ceteris paribus* worse than allowing harm. Nevertheless, philosophers have found the analysis of the difference elusive and the ensuing controversy has led many to skepticism about the coherence of the distinction and/or its moral significance.

Our goal here is to say “What’s the Dif?”

We take this analytical job to be the starting place for a theory of *agent* or, more generally, *object causation*. We invoke this sort of causation whenever we speak of a person or an object or a kind of thing as a cause; and we do so all the time—when we say that the moon causes the tides but not lunacy, that Vesuvius destroyed Pompeii, that the rhinovirus causes colds, that Oswald killed Kennedy. Such talk is vitally important; as when we ask what virus causes a disease, or which person is the murderer, or who let the dogs out. And it is in this idiom that we invoke the Dif: the secret service allowed, but did not cause Kennedy’s death; the Federal Reserve did not prevent the recession, but it did not make it happen; by failing to destroy the rhinovirus, the immune system allows us to catch cold, but the common cold is not an autoimmune disease.

Of course, objects enter into causal relations only by doing things—that is, by participating in events—which have causal upshots. But we need a separate account of object causation because the causal consequences of what an object or agent does are not always things that the *agent* or *object* causes. That is, we do not describe an object, *o*, as a cause of some outcome *merely* because *o* does something that is a cause of that outcome. A fire alarm’s failing to go off can be a cause of the destruction of a building but there is nevertheless a difference between the defective fire alarm that fails to sound when a fire breaks out and the defective fire alarm that short-circuits and

\* We thank Jonathan Bennett, David Braybrooke, Steve Finlay, Matt Hanser, Mark Heller, and Jonathan Schaffer for their helpful comments on earlier versions of this paper. We also thank audiences at the Southern California Philosophy Conference and at the Philosophy Department of the University of California, Santa Barbara.

causes the fire. It is the same kind of difference as between a lazy night watchman, liable to sleep through a fire, and a malicious watchman bent on arson. One object allows; the other causes the bad outcome.

In order either to cause or allow an outcome, something that the agent does must be a cause of the outcome. Agents cause outcomes by doing things. These doings, also called “actions,” are causal *relata*; they are events in which the agent figures in the way that objects figure in events.<sup>1</sup> The ontology of actions is no more or less mysterious than the ontology of events generally, and while that might be very mysterious indeed, we think that nothing in our account trades upon or compounds those mysteries.<sup>2</sup>

Able pushes Vic into the lake. Though he could, Baker does not throw Vic a life preserver. Vic drowns. Able does something that causes Vic’s death and so does Baker. Baker’s refraining from throwing a life preserver is something Baker does. We see no reason not to call it an action nor do we see any reason to deny that this action is a cause of Vic’s death.

Had Baker thrown the life preserver, Vic would not have died. Able causes Vic’s death. But while it is correct to say that *something that Baker does* is a cause of Vic’s death, we do not think it right to say that *Baker* causes Vic’s death.

The difference between Able and Baker is a matter of the *way* they do the things that cause death. The *way* in which someone does something is a matter of *how* they do it.

Questions about how things are done are answered using the ‘by’-locution.<sup>3</sup> We say “A does X by doing Y.” The analysis of ‘by’ statements has been the subject of much heated controversy.<sup>4</sup> We are not

<sup>1</sup> Our account of the Dif could be readily recast in the language of fact causation, however.

<sup>2</sup> Nothing in our account will turn on any special assumptions about the nature of events or actions. There is nothing up our metaphysical sleeves. In particular we do not frontload our understanding of event causation in some way which excludes omissions or background conditions as causes. Neither do we assume any distinction between causally necessary conditions and “the” cause of an outcome. We take a maximally liberal construal of what count as actions and events. Thus we count the striking of a match as an action/event and a cause of its lighting but would also count the match’s being dry as an event and a cause and the agent’s refraining from getting the match wet as an action and a cause. In this we part company from those who would seek to ground the Dif in the fundamental ontology of events or the theory of event causation. Cf. Judith Jarvis Thomson, “Causation: Omissions,” *Philosophy and Phenomenological Research*, LXVI (2003): 81–103.

<sup>3</sup> The “‘by’-locution” locution is owed to Jonathan Bennett, *The Act Itself* (New York: Oxford, 1995).

<sup>4</sup> For an extended discussion of ‘by’ see Bennett, *Events and Their Names* (New York: Oxford 1988), chapters 12–14, and *The Act Itself*, chapter 2.

interested here in entering this fray. Any plausible analysis of ‘by’ talk must respect common sense and common usage and any such analysis will, we believe, sustain our analysis of the Dif. We assume only this much, and it is an assumption shared by every analysis of the ‘by’ relation of which we are aware.

*A* does *X* by doing *Y* only if:

- (i) *A* does *X*; and
- (ii) *A* does *Y*; and
- (iii) (ii) is an Insufficient but Necessary part of a condition which is logically, nomologically, or causally Sufficient for (i).

The answer, when there is one, to the question “How did *A* do *X*?” will take the form “by doing *Y*” and it will always be on the cards to ask, in turn, “How did *A* do *Y*?” and so on. *Ad infinitum*? No. It seems, typically at least, that these ‘by’ chains come to an end when they arrive at what some call “basic actions.” Basic actions are doings for which there is no ampliative answer to the question “How did *A* do it?” Thus, we might say that someone caused the wreck by signaling and signaled by raising his hand, but we may have no answer to “How did he raise his hand?,” no way of filling out, “he raised his hand by. . . .” “He just did it” we are wont to say.<sup>5</sup>

Whether or not these ‘by’ chains must *always* come to an end is a question we shall not address here, but whether or not they do, we may say that the whole story about the *way* in which someone performs an action is a matter of the whole story about how they do what they do as revealed by these chains of ‘by’ dependence. So if Able and Baker both do *X* and both do *X* by doing *Y*. . . they may still do *X* in different ways if one does *Y* by doing *Z* and the other does not.

Sometimes ‘by’ relates a *causally* INS condition. Thus:

- (1) He killed him by pushing him in the lake.
- (2) She put out the fire by dousing him with water.
- (3) He rang the bell by pushing the button.

But not all ‘by’s are causal. Consider:

- (4) He broke his promise by not writing.
- (5) She made it unanimous by voting “aye.”

In cases like these the unstated conditions will complete a condition which is *logically* sufficient for the outcome: he had promised to write and he did not, she voted aye and so did everyone else.

<sup>5</sup> Though we might be able to imagine circumstances in which someone raised his hand by doing something else. Cf. Bennett, *The Act Itself*, p. 37.

There is another sort of noncausal ‘by’ that will be important to what follows. Consider the question “How did Baker fail to throw the life jacket?” The answer may be important. If Baker tried to throw the jacket but failed by slipping on the wet dock his culpability might be diminished. Suppose that, in our case, the answer is:

(6) Baker failed to throw the life jacket by standing on the dock, twiddling his thumbs.

Here the unstated conditions would include additional facts about Baker’s behavior and environment—he did not do anything tricky with his feet, there were no mechanisms in his surround to convert thumb twiddling into life jacket throwing—facts that, together with the laws of nature, entail that Baker failed to throw the life jacket. Which is to say that if (6) is true, it will be because Baker’s thumb twiddling is an insufficient but necessary part of a *nomologically sufficient* condition for Baker’s failure to throw the life jacket. Even so, (6) is not causal. Though it is nomologically INS, the thumb twiddling is not a cause of Baker’s failure. Had Baker not twiddled his thumbs, the jacket might still not have been thrown. He might have put his thumbs in his ears, or dealt cards, or walked away . . . and the failure would still have occurred.<sup>6</sup>

The whole story about how someone performs an action will be told by chains of ‘by’ statements and the whole story might be causal. Thus it might be that *A* does *X* by doing *Y*, where doing *Y* is a cause of *X*, and that he does *Y* by doing something that causes *Y* and so on, *ad infinitum*, or until we arrive at basic actions.

In this sort of case the ‘by’-chain is just a causal chain and, insofar as causation is transitive,<sup>7</sup> each link in the chain will be a cause of *X* and of the outcome *X* causes. Interestingly, this can be so even if each link in the ‘by’-chain is not a cause of the preceding one. Thus: Swift won the medal. How? By breaking the record. How? By running a three minute mile. How? By sprinting for the last two minutes. The connection between running the three minute mile and breaking the record is logical, not causal. That he ran the race in three minutes, together with the fact that the preceding record was four minutes,

<sup>6</sup> Nomologically INS conditions that are not causal are the bane of attempts to analyze causation in terms of laws. Notoriously, the barometer’s fall is INS for the thunderstorm, but does not cause it.

<sup>7</sup> Note, however, that our account does not rely on the assumption that causation is transitive. For a discussion of arguments against the claim that causation is transitive, see Ned Hall, “Causation and the Price of Transitivity,” this JOURNAL, xcvi, 4 (April 2000): 198–222.

entails that the record was broken. And we might have similar qualms about saying that his two minute sprint is a cause of his three minute race. The relation here, we might want to say, is that of part to whole, not cause and effect. Be that as it may, note that each link of this chain—the record breaking, the three minute race, the two minute sprint—can be properly described as a cause of Swift’s winning the medal, even if it does not cause the doing that immediately precedes it in the chain.

In this sort of case and in the case where the ‘by’-chain is a simple causal chain, each link in the chain of answers to the question “How did *A* do *X*?” reveals a cause of *X*. Let us call these ways of doing things—where causation reaches all the way down—*wholly causal* or *causal* for short.

Not every way of doing things is causal. Sometimes when we fill out the answer to the question “How did *A* do *X*?” we arrive at something *A* does that is not a cause of *X*. As we have just observed, the whole story about Baker’s role in Vic’s death turns noncausal when it gets to (6). The thumb twiddling and the failing to throw the life jacket are both things that Baker does. He does the latter by doing the former. But the thumb twiddling does not cause Vic’s death. And while we could go on to ask “How does he twiddle his thumbs?” the answers we get will not be causes of Vic’s death.<sup>8</sup>

In this sort of case, the explanation of how *A* does *X* runs to doings that are not themselves causes of *X*. Let us call cases like these—where the causation does not go all the way down—*noncausal*.

The difference between an agent’s doing something in a causal as opposed to a non-causal way is, we maintain, precisely the difference between an agent’s causing something as opposed to allowing it to happen.<sup>9</sup> Assuming that, in the final ‘by’ analysis, agents do everything they do by performing basic actions, we could say:

<sup>8</sup> Some folk have complained that it feels odd to ask “How did he do *X*?” when *X* is an omission. We agree. The actions folk call omissions are, we claim, precisely the ones that are done noncausally. Folk rarely bother to ask how omissions are performed because they recognize that the answer is not going to reveal a cause and causes are usually what folk are interested in.

<sup>9</sup> While we will sometimes speak of “*the way*” an agent does *X*, we do not deny that, for any *X*, there may be many different legitimate ‘by’ stories one can tell. One might say, with equal truth, “He bought the house by cashing in his life savings” or “by submitting a higher bid” or “by duping the appraiser.” Each may be true and each begins a ‘by’ chain which will comprise different actions on anyone’s account. We conjecture that if any of these stories are causal in our sense then any other legitimate ‘by’ story about how the agent does *X* must be causal as well. But this conjecture is no part of our account of the Dif. On our account an agent causes *X* if there is *a* causal ‘by’ story about how he brings about *X*.

(7) An agent causes an outcome *iff* the outcome is caused by a basic action of the agent.

Without appealing to the idea of basic actions, we can say:

(8) An agent *A* causes an outcome *O* *iff* *A* does something, *X*, that causes *O* and *A* does *X* causally.

Take another case. A runaway train is hurtling down the track towards an unsuspecting group of cub scouts. A bystander sees that if she throws the switch the train will be diverted to another track, where only one person, Joe, will be killed. Throwing the switch will cause Joe to die. Refraining from throwing the switch will cause the cub scouts to die. Whatever the bystander does, *something she does* will be a cause of the death of someone. But that is not to say that the *bystander* will cause a death no matter what. There is no way for the bystander to throw the switch except causally. So if she throws the switch, she will cause Joe's death. She will also cause the cub scouts not to die; that is, she will prevent their death. On the other hand, there is no way for the bystander to refrain except noncausally. So, by refraining, she will not have caused the death of the cub scouts. Thus: if the bystander refrains from throwing the switch by remaining perfectly still, it will nevertheless be false that her refraining still causes the scouts' dying. She could have refrained without remaining perfectly still, and the cubs would still have died. Likewise, while the bystander's refraining from throwing the switch is a cause of Joe's not dying, our analysis does not require us to say that in that case the *bystander* would have caused Joe not to die. More colloquially, we need not praise the bystander for preventing Joe's death but can say, instead, that she only allowed him to live.

We have defined the difference between causing and allowing in causal terms but in many of our examples we have used counterfactual dependence as a test of causal connectedness. We have taken the fact that an outcome might have happened even if the agent had not done *X* as evidence that doing *X* was not a cause of the outcome. But counterfactual dependence is not an infallible guide to causal dependence. Notoriously it fails in cases of causal preemption. Here is an example due to Frances Howard-Snyder.<sup>10</sup>

An assassin, A. Sassan, is preparing to assassinate Victor by shooting him.  
A second assassin, Baxter, is waiting across the street watching Sassan

<sup>10</sup> "Doing vs. Allowing," in Edward N. Zalta, ed., *Stanford Encyclopedia of Philosophy* (Summer 2002 Edition on the internet).

to ensure his success. If Sassan shows any signs of hesitation, Baxter will shoot Victor himself. Suppose Sassan knows about Baxter and his intentions and also knows that he can turn his gun on Baxter instead of on Victor if he so chooses. Although this thought crosses his mind, he quickly suppresses it, since he is committed to Victor's annihilation. He shoots Victor and Victor dies instantly.

Here it is false that if Sassan had not shot, Victor would not have died. But surely, Howard-Snyder insists, Sassan caused Victor's death by shooting. We agree. The case is one of preemption and in such cases causal dependence does not pair with counterfactual dependence. Victor's death is caused by Sassan's shooting, even if it does not counterfactually depend on it. That is why cases of preemption are standard counterexamples to the thesis that causal dependence is simply counterfactual dependence.

Even so, given our reliance on causation, it might be held that we have a special obligation in such cases to say why the dependence is causal when it is not counterfactual. While it is not our business here to endorse an analysis of event causation it seems to us that David Lewis has the right answer.<sup>11</sup> Roughly, Lewis proposes that to say that  $c$  causes  $e$  is not to say not that  $\sim c > \sim e$ , but rather that the timing and manner of  $e$ 's occurrence (or the occurrence of an event like  $e$ ) is counterfactually dependent on the timing and manner of  $c$ 's (or an event like  $c$ 's) occurrence. Lewis calls this complex relation "influence" and it may be that  $c$  influences, and hence causes,  $e$  even if  $\sim c > \sim e$  is false. Thus, in the example above, while it is false that Victor would not have died if Sassan had not shot, it is true that the manner of Victor's death is highly dependent on when and how Sassan shoots. If Sassan had shot sooner or had shot differently (aiming higher or lower), then Victor would have died in a different way at a different time.

This seems to us right or at least on the right track and it underlines a phenomenon illustrated in the examples of allowing above. In non-causal cases, the way in which someone performs an allowing is usually only loosely connected with the way they move their body. There are many ways to skin a cat, but there are many, many more ways to allow a cat to be skinned.

We think that something like this observation underlies Jonathan Bennett's analysis of the Dif. Bennett draws a distinction between "making" and "allowing" in terms of an agent's alternative actions. An agent moves his body and an outcome occurs. Had the agent

<sup>11</sup> "Causation as Influence," this JOURNAL, xcvi, 4 (April 2000): 182–97.

moved his body in a different way, the outcome would not have occurred. But now suppose that only a small fraction of the agent's range of alternative actions would have made a difference to the outcome. Had he moved in almost any other way, the outcome would still have occurred. In that case, Bennett says, the agent allowed the outcome. On the other hand, suppose the outcome would not have happened if the agent had moved in any of most of the alternative ways he could have moved. In that case, Bennett says, the agent causes the outcome. Thus, by standing by, the bystander allows the cub scouts to die because most of the alternative ways she could have moved would have the scouts dying anyway. On the other hand, had she thrown the switch, the bystander would have caused Joe's death, since most of the alternative ways she could have moved would have involved no switch throwing and no death.

Bennett's analysis is certainly the most successful and best argued for account heretofore. Lewis's analysis of causation as influence is useful because it can help us understand something that needs explaining: why Bennett's analysis works as well as it does over a large range of cases in capturing our intuitions about the Dif. This needs explanation because Bennett's analyses are neither obvious nor obviously connected to the moral distinction the Dif is supposed to make. As Bennett observes:

The...distinction that I have defined *obviously* has no basic moral significance: if someone moves in a way that causes or makes probable some bad upshot, nobody would think that the moral status of his conduct depends on how many other movements by him would have done the same.<sup>12</sup>

This leads Bennett to conclude ultimately that the Dif makes no real moral difference.

This seems to us very wrong. Bennett's account works, when it works, because his criterion measures the extent to which the *way* someone moves his body *influences*, in Lewis's sense, whether an outcome occurs. What is really being measured is whether the agent's *body movements* cause the outcome. Thus the bystander's failure to throw the switch is a cause of the cub scouts' death. She fails to throw the switch by standing perfectly still; the fact that she could have moved in many other ways without preventing their death is relevant

<sup>12</sup> *The Act Itself*, pp. 102–103; emphasis in original.



because it means that the way she actually moves her body is not a cause of the scouts' death. In turn, it seems to us that whether or not someone's body movements cause an outcome is *obviously* relevant to the moral status of her conduct.

Bennett's measure does not perfectly track influence. Howard-Snyder's Sassan case was offered as a counterexample to Bennett's analysis and in this it seems to us wholly successful. On the other hand, there is another sort of counterexample that is sometimes brought against Bennett that might also count against our own analysis. Here is Bennett's telling of the story, which we will call *Immobility*:

Henry is in a sealed room where there is a fine metallic dust suspended in the air. If he keeps stock still for two minutes, some dust will settle in such a way as to close a tiny electric circuit which will lead to some notable upshot *U*. Thus, any movement from Henry and *U* will not obtain; perfect immobility, and we shall get *U*.<sup>13</sup>

Henry does not move. Bennett's analysis rules that, by remaining perfectly immobile, Henry makes *U* happen. Some find this counterintuitive. Bennett thinks the intuition can be explained away by the fact that in real life situations total immobility is rarely productive of the outcomes in which we have a quotidian interest.

What does our account say about the immobility case? The answer depends upon a prior question about the causal connection between events. Should we say that:

(9) Henry's immobility causes the dust to settle.

We are inclined to so say, but not everyone would agree. Some philosophers want, for various reasons, to deny that "omissions" or "inactions" can be causes.<sup>14</sup> On such views of *event* causation, (9) might be held to be false. It would follow, on our account of *agent* causation, that Henry does not cause either the settling of the dust or *U*.

In this way, the case illustrates how our theory of agent causation can serve competing intuitions about event causation. But we are suspicious of accounts that try to count out omissions as causes. We suspect that they are born of the desire to serve, at the level of event causation, intuitions that properly belong to the theory of agent causation. Indeed, we think that the failure to distinguish between agent and event causation—the failure to distinguish between what

<sup>13</sup> *The Act Itself*, pp. 97–98.

<sup>14</sup> For a succinct overview of these views of causation see Schaffer, "The Metaphysics of Causation," in *The Stanford Encyclopedia of Philosophy* (Spring 2003 Edition).

agents cause and what is caused by the things they do—is the source of much confusion in the metaphysics of causation.

So we agree with Bennett's verdict in this case. Nevertheless, we think his account fails and fails not just because of preemption counterexamples. Consider another case of our own:

*Mobility:* Henry is in a sealed room with an extremely sensitive motion detector. If Henry moves at all, the detector will trigger some notable upshot *U*. Only if he keeps stock still for two minutes will *U* fail to obtain. Henry moves.

On Bennett's analysis, if Henry moves, he will only have allowed *U* to occur since *U* would have happened for almost every way Henry could have moved. This seems to us obviously false. Henry's moving causes *U* by as clear a chain of cause and effect as a carom of billiard balls—never mind that he could cause the outcome by moving in any way at all. So Henry causes *U*.

Causation makes the Dif, and, it seems to us, it is a difference which manifestly bears all the moral significance which tradition and commonsense confer on it.

## II

Let us consider some cases.

First, as points of moral and causal reference, we have paradigm cases of agent causing:

*Push:* A cart stands at the top of a hill. Agent pushes it. The cart rolls down the hill and fatally injures a child.

and allowing:

*Stayback:* The cart is already rolling; Agent could but does not interpose a rock which would stop it. The cart rolls down the hill and fatally injures a child.<sup>15</sup>

Any reader who does not feel a *prima facie* moral difference in these cases probably need not read on. For those who intuit a difference, however much they might suspect or disdain that intuition, let us use these cases as touchstones on the rocky road ahead.

Things quickly get tendentious when we turn to:

<sup>15</sup> The cart setup and the specific examples Push, Stayback, and Kick are due to Bennett, *The Act Itself*, p. 67.

*Dislodge*: A cart stands at the top of a hill, its wheel chocked by a rock. Agent kicks the rock away. The cart rolls down the hill and fatally injures a child.

*Kick*: The cart is rolling towards a point where there is a rock that would bring it to a halt. Agent kicks away the rock; the cart rolls down the hill and fatally injures a child.

Do the kicks in *Dislodge* and *Kick* count as causes of death? There are theories of causation that would say not. *Dislodge* is a case of what Jonathan Schaffer calls “causation by disconnection” and cases of this sort are counterexamples for theories that require some flow of force or energy from cause to effect.<sup>16</sup> *Kick* is an instance of the broader class of cases which the literature calls “preventing a preventer” and they will seem problematic for any theory that requires chains of spatiotemporal contiguity or persisting processes between cause and effect or any theory that wants to prohibit absences (such as, of the obstructing rock), or omissions (such as, the cart’s not striking the rock) as causes.<sup>17</sup>

In respect to the purely causal issue, it seems to us that any theory that fails to count the kicks in *Dislodge* and *Kick* as *bona fide* causes is *prima facie* dubious. As others have pointed out, causation by disconnection, by the prevention of preventers, by omission and absence, are ubiquitous in everyday life and are there unhesitatingly counted as real causation by anyone who does not have a metaphysical axe to grind.<sup>18</sup>

On the other hand, with respect to the division of cases into doings and allowings, things are not so clear. There is a temptation, stronger in *Kick* than *Dislodge* but undeniably present in both, to describe these agents as having, by their actions, “allowed the cart to roll down the hill and kill the child.” This way of speaking, in turn, might incline us to a theory of the DIF which pushes disconnection and preventer prevention to the allowing side of the divide.

<sup>16</sup> See, for instance, David Fair, “Causation and the Flow of Energy,” *Erkenntnis*, xiv (1979): 219–50.

<sup>17</sup> Cf. Phil Dowe, *Physical Causation* (New York: Cambridge, 2000); David Armstrong, “Going Through the Open Door Again: Counterfactual versus Singularist Theories of Causation” in Gerhard Preyer and Frank Siebelt, eds., *Reality and Humean Supervenience* (Lanham, MD: Rowman and Littlefield, 2001), pp. 163–76; Helen Beebe, “Causing and Nothingness,” in John Collins, Ned Hall, and L.A. Paul, eds., *Causation and Counterfactuals* (Cambridge: MIT, 2004), pp. 291–308.

<sup>18</sup> See Schaffer, “Causation by Disconnection,” *Philosophy of Science*, LXVII (2000): 285–300; Lewis, “Causation,” in *Philosophical Papers*, Volume 2 (New York: Oxford, 1986), pp. 159–213; H.L.A. Hart and Tony Honore, *Causation in the Law* (New York, Oxford, 1959).

The problem with this route is that the differences between preventing and enabling, acting and omitting seem as evanescent as the Dif itself. Here is Bennett on the proposal to treat Kick as an allowing because the agent only prevents a preventer:

But then in Push, when the vehicle was standing motionless, that was because it was prevented from rolling by small stones, holes in the ground, wet grass, and so on; so an agent in pushing the vehicle was preventing these things from preventing it from moving; yet in this case we do not say that in pushing the vehicle he allowed it to move. What makes the difference? I am not sure; though it seems to involve the difference between removing obstacles (Kick) and enabling something to overcome obstacles (Push). Enough of this! We should not look for help from the meaning of 'allow' which is a complex shallow mess.<sup>19</sup>

Bennett's position is that Kick belongs, where his account would place it, on the making (Push) side of the Dif.

We have already committed ourselves to counting the kicks in Dislodge and Kick as causes of the cart's motion and that commits us to treating these agents as causing the child's death. In support of this way of carving things up, we may note that whatever temptation there may be to use the language of 'allowing' in these cases there is no corresponding temptation to treat these agents *morally* on a par with the agent of Stayback. In real life cases of disconnection and preventer prevention—the hangman springing the gallows' trap, the vandal who removes the guard rail at the tourist lookout or the greedy pharmacist who waters the vaccine—we count the agents as killers on a par with Push and not bystanders like Stayback.<sup>20</sup>

Let us acknowledge the obligation to explain away the contrary intuition here—to do something to clean up the mess about which Bennett complains—and take up a more pressing problem.

The problem is *Step Away*.

*Step Away*: The cart is already rolling; Agent is in its path at the top of the hill and if he remained where he was the cart would bump into him harmlessly and stop. Agent realizes this but steps out of the way and the cart rolls down the hill and fatally injures a child.

It seems that the only difference between Step Away and Kick is

<sup>19</sup> *The Act Itself*, p. 68.

<sup>20</sup> It is easy to get derailed by features of examples that seem trivial. One's first intuition is that the agent of Push is more likely to be counted guilty of murder in the first degree than the agent of Kick but that is only because, in the normal course of things, kicking away is not a likely technique for premeditated murder.

that in the former case it is the agent's body, not the rock, which is the obstacle to the cart. But this seems not much of a difference. Indeed, we can narrow the difference more by imagining that in Step Away the agent steps away by kicking a rock by the side of the path, so that the agent's body moves precisely as it does in Kick and, of course, with the same upshot—the death of the child.

This argues for treating Step Away as a killing on a par with Kick. The problem is that Step Away looks like an instance of *ducking*.

*Duck*: Agent sees that the bad guy is about to shoot in his direction.  
Agent ducks and the shot kills a child.

Ducking was introduced to the literature by Christopher Boorse and Roy Sorensen.<sup>21</sup> They contrasted ducking with *sacrificing*.

*Sacrifice*: Agent sees that the bad guy is about to shoot in his direction.  
Agent grabs a child and interposes its body between himself and the gunman. The bad guy shoots and the child is killed.

Boorse and Sorensen thought it clear that the agent of Duck is no killer, while the agent of Sacrifice is. If we agree with that, then it seems that we ought to treat Step Away as a case of ducking and we are stuck with explaining the difference between Kick and Step Away.

Bennett's account offers a way to do so. Plausibly, of all the ways of moving available to the Kick agent, only a few constitute a kicking of the rock. On the other hand we can imagine that there are many different ways that the Step Away agent might move, any one of which would be a stepping (or hopping or jumping . . .) away. This would allow Bennett to treat Step Away as an allowing and Kick as a killing.

The trouble with this is that we can also imagine that there are *not* a lot of alternative ways of stepping away. Jeff McMahan<sup>22</sup> describes this case:

*Lunge Away*: Like Step Away except that Agent can only remove himself from the path of the cart by hastily executing a precise set of specific body movements.

Bennett must count this as a killing and of course will have to precisely parse the continuum of cases where varying restrictions on range of motion transform a Step Away agent's *allowing* to a

<sup>21</sup> "Ducking Harm," this JOURNAL, LXXXV, 3 (March 1988): 115–34.

<sup>22</sup> "A Challenge to Common Sense Morality," *Ethics*, CVIII (1998): 394–418, see p. 404.

Lunge Away agent's *killing*. And if this seems arbitrary, Bennett can always reply that that is why he thinks it an illusion that *the Dif* morally matters.

One begins to see why Boorse and Sorensen introduced making sense of ducking and sacrificing cases as "an adequacy condition on solutions to the act/omission problem". Interestingly, though, they did not think that the Ducking-Sacrificing divide is equivalent to the Doing-Allowing Dif. Their evidence was a case we can call *Drive Away*:

*Drive Away*: Agent's car is parked between a rolling cart and a child. Agent sees the child in the cart's path but drives off to avoid an ugly dent.

They argued that if Duck is an allowing then so is Drive Away. But, they insisted, Drive Away, unlike Duck, is a case of "culpable homicide." The latter agent ". . . can be convicted of murder or manslaughter in any American court. On the other hand, had [Agent] merely failed to interpose his car between rock and child, he would be innocent in both Anglo-American law and, presumably, an ethics that takes the act/omission distinction seriously."<sup>23</sup>

Considering this case, Alison McIntyre concludes that since the *only* difference between Drive Away and Duck is a moral one this shows that our apparently causal judgments are sometimes fundamentally controlled by our moral assumptions about what one is obliged to do:

Since the cases are similar except in their moral characteristics, what lies behind this judgment must be the assumption that there is a moral duty to risk damage to your car to protect the child, and a consequent expectation that you will do so but that there is no duty for you to risk injury to yourself—hence, moving your body out of the way does not cause the child to be injured and is not a case of causing the child's death.<sup>24</sup>

But it would be very remarkable indeed if our *moral* assumptions were to drive our *causal* judgments in these cases. Consider that, notwithstanding the expectation that one need not sacrifice oneself, we are not in the least tempted to deny that the agent who shoots an assailant in self defense *causes* the death of his assailant.

Shifting perspective, McMahan offers another sort of case that

<sup>23</sup> Boorse and Sorensen, p. 127.

<sup>24</sup> "Guilty Bystanders? On the Legitimacy of Duty to Rescue Statutes," *Philosophy and Public Affairs*, xxiii (1994): 169–70.

suggests that our intuitions on *any* of these cases might be altered simply by changing its surround:

*Interpose Then Kick:* The cart is already rolling. Agent places a rock in the path of the cart. If left in place, the rock would stop the cart; but before the rock can stop the cart, Agent reconsiders and kicks the rock away. The cart rolls down the hill and fatally injures a child.

McMahan argues that even if Kick is killing, Interpose Then Kick looks more like allowing.<sup>25</sup> After all, if we call this killing what should we say about *Walk By*?

*Walk By:* The cart is rolling down the hill. Agent walks into its path. If he remained where he is, the cart would bump harmlessly into Agent and stop. But before the cart reaches him, Agent walks on.

And how, in turn, should we distinguish Walk By from Step Away?

Interpose Then Kick and Walk By are only two of a whole genre of cases that McMahan dubs “withdrawing aid.”<sup>26</sup> Other examples:

*Dutch Boy:* Dutch Boy spots a leak in the dyke. The leak threatens to ultimately wear away at the structure, collapse the dyke, and destroy the village. Dutch Boy plugs the leak with his finger but, after a time, tires of his situation and pulls his finger out. The leak resumes, destroys the dyke, and wreaks havoc on the village.

*Involuntary Donor:* Agent finds himself connected by a medical apparatus to a sickly violinist. The apparatus enables Agent’s body to sustain the violinist. Tiring of his situation, Agent unhooks himself from the apparatus. The violinist dies.

Withdrawing aid, McMahan argued, is so different from either killing or allowing as to require an entirely separate status.

So what do *we* say about all these cases? Which are killings, which are lettings die? A first answer: in some cases, both.

Traditionally, accounts of the Dif have been *summary*, in the sense that taking everything the agent does into account they judge that he has either caused or allowed an outcome. This is not so with our account; an agent may cause a death by doing one thing and allow it by doing another. Able caused Vic’s death by pushing him

<sup>25</sup> “This background detail inclines us to locate this kick on the negative side of the generic positive-negative divide”—McMahan, “A Challenge to Common Sense Morality,” p. 402.

<sup>26</sup> McMahan, “Killing, Letting Die, and Withdrawing Aid,” *Ethics*, CIII (January 1993): 250–79.

in the lake and allowed him to drown by failing to throw him a life jacket.

Many of the cases just surveyed are bivalent in this way. Thus, after the Dutch Boy has removed his finger it is easy to imagine that he could reinsert it and save the day. So we should say that by failing to put his finger back in the Dutch Boy allows the ruin of the village even if we also say that he causes it by pulling his finger out in the first place. Likewise we could imagine that the agents in Involuntary Donor and Step Away might have had second thoughts. The Involuntary Donor could reconnect the machine. The agent in Step Away could step back. By failing to repent, these agents allow the harmful outcomes they cause.

That one can allow as well as cause explains our temptation in cases like Kick to describe the agent as “allowing the cart to roll down the hill” since that is precisely what the agent does in the interval before and after he kicks the stone away. If the interval of allowing is short, as it will typically be in a case of Dislodge, there is less temptation to use ‘allowing’ talk. And where little or no allowing is likely, as in Push, there is little or no temptation to describe what has gone on as “allowing the cart to roll.”<sup>27</sup>

More careful story telling can remove the dissonance. Suppose that the damn bursts the instant the finger is withdrawn; that the patient dies the moment the machine is disconnected; imagine that the stone cannot be reinterposed. Thus redescribed, we think the temptation to see these as allowings recedes.

So, setting aside the distraction that some of these cases may be allowings as well as killings, which ones are also killings?

Let us put our cards on the table. We think that with the exception of Stayback, *all* agents in the cases we have considered in this section—Push, Dislodge, Kick, Step Away, Duck, Sacrifice, Lunge Away, Drive Away, Interpose Then Kick, Walk By, Dutch Boy, Involuntary Donor—are killers; each causes, not merely allows, death.

Let us quickly add, however, that we do *not* hold that these cases are *morally* equivalent. Indeed, we offer it as a central argument for our account of the Dif that it can make sense of the moral *differences* among these cases.

### III

The Oxford English Dictionary defines consequentialism as “an ethical doctrine which holds that the morality of an action is to be judged

<sup>27</sup> We should note, but will not here elaborate upon the other complication here; namely, that as an upshot of the agent’s actions the *rock* allows, where it would otherwise have prevented, the cart to roll down the hill.



solely by its consequences.” But, of course, it is not so. The death of an innocent is a bad consequence, but no consequentialist would condemn an action just because it has this upshot. It depends on the alternatives. If the agent’s choice is between killing one innocent and killing many, the consequentialist would recommend killing one. A consequentialist wants to say things like “it is better to kill one innocent than two,” but, of course what this means is that you should kill one rather than two *if those are your alternatives*. It does not mean that if Able kills one innocent and Baker kills two then Able’s act must be better than Baker’s. It depends on the alternatives. Better to say, “other things being equal, one killing is better than two,” where what are equal are the *moral* consequences of the alternatives to the killings.

A deontologist is someone who thinks that *how* an action is performed is morally significant. The deontological commonplace with which we began was that it is worse to kill someone than to allow someone to die. But, come to think of it, that cannot be the assumption which has been animating our intuitions about these cases. Taken at face value, that principle would recommend allowing someone to die over killing them. Given such a choice this might or might not be sound advice, but that is *not* the sort of case we have been considering.

Our paradigmatic allower, Stayback, does not choose between allowing the child to die and killing him. By staying back he allows the child to die rather than *preventing* his death. And our paradigm killer, Push, has not chosen to kill the child rather than allow him to die; he has chosen to kill over allowing the child to *live*.

Nor does it help to qualify by saying “other things being equal it is worse to kill than allow.” For what is being held equal in our judgment that Push is worse than Stayback? One might think the alternatives are morally equivalent since, in both, a child does not die. But that would be to judge the alternatives by their *consequences*. The alternative *actions* in these cases are allowing a child to live and saving a child’s life and the same deontological intuitions that tell us that *causing* is worse than *allowing* the death of a child also cry out that *saving* a child’s life is *better* than *allowing* it to live. If the agent of Stayback had not stayed back we would have praised him, but we do not praise people for refraining from killing children.

So the principle *really* at work in our judgment that Push is worse than Stayback seems to be:

Killing someone, when the alternative is to allow that person to live, is worse than allowing someone to die, when the alternative is to prevent that person’s death.

Now all of this might not be worth saying if the alternatives were

always the same. But are they? Must it always be that the alternative to killing is to allow life to continue, that the alternative to allowing a death is saving a life? More generally, must it always be that an agent's only alternative to causing an outcome will be to allow it not to happen?

The answer will depend on one's theory of the difference between doing and allowing. On our account, the answer is "no."

To see this, consider the difference between Kick and Step Away. In both of these cases a child dies as a causal upshot of an agent's basic action. In one case it is a kick; in the other, a step. As we observed above, the choreography of these actions may be precisely identical. On our account, both agents cause a child's death. But there is a difference.

Consider the situation if the agent of Kick had not kicked away the rock. His refraining from kicking the rock would have been a cause of the child's survival. How might he have refrained? Well, by stepping back, or dancing a two-step, or sitting down.... It does not much matter. Refraining from kicking a rock is no basic action; to perform it the agent would have to do some action or other (other than kick that is). Any action that, in the circumstances, nomologically suffices for the rock not moving would serve. But nothing that the agent does in this vein will be a *cause* of the child's survival. This is to say, if the agent of Kick had not kicked the rock away he would only have allowed the child to live; he would not have saved the child's life.

Compare this to the case of Step Away. Suppose that agent had not stepped away; that he had, instead, remained in place. How might he have done that? Well, he might have stepped briefly to the side then back again. Or he might have lunged out of the way and then crawled back into the cart's path. Or the agent might have performed that most basic of basic actions: he might have just stood still. In any of these alternatives, the cart would be stopped by a basic action of the agent. So, by our reckoning, the agent would have caused the child not to die, or in other words, prevented the child's death.

This, in turn, is why, had the Step Away agent not stepped away, we would praise him for saving the child's life. If the bump he endured from stopping the cart were great, we would praise him all the more. On the other hand, if the Kick agent had just stayed back he would earn no praise. He would only have allowed the child to live. In so doing he would have obeyed morality's minimal requirements but no hero he.

The Kick agent chooses to cause the child's death rather than allow it to live. The Step Away agent chooses to cause the child's death rather than prevent its death.

It is this difference in the moral weight of the *alternative* actions open to them that explains, we think, the moral differences between the agents with respect to the actions they *actually* take. Both agents cause a child's death, but the agent of Step Away is morally on a par with the agent of Stayback, who only allows a death, while Kick looks every bit as bad as Push.

Generalizing this observation leads us to what we shall call the *Deontological Equivalence Principle*. The principle might be put like this:

*Deontological Equivalence*: Causing a harm, when the alternative is preventing that harm, is morally equivalent to allowing a harm when the alternative is preventing that harm.

Where the harm is death:

Causing a person's death, when the alternative is preventing that person's death, is morally equivalent to allowing a person to die, when the alternative is preventing that death.

More colloquially:

Causing a death you would otherwise be preventing is no worse than allowing it.

On this principle, the agents of Step Away, Duck, Lunge Away, Walk By, Dutch Boy, and Involuntary Donor, though they all cause death, are morally no worse than Stayback; no worse, that is, than the person who allows someone to die rather than save him.

The agents in these cases cause death by refraining from preventing it. An ordinary killer of the Push or Kick sort causes death by refraining from allowing the victim to live. The Equivalence Principle allows us to treat an agent as causing a death without having to regard him as morally on a par with an ordinary killer. And separating the moral from the causal judgments allows us accommodate our intuitions about the differences in these cases without having to gerrymander our account of the Dif.

To understand the difference between our account and mere gerrymandering it is important to understand what we are *not* arguing when we advocate the Equivalence Principle.

Looking at these cases, one might think that what they have in common is, roughly speaking, that the benefits that the agents deny victims are ones that the agents themselves have conferred. Thus, the barriers to harm that Dutch Boy and Involuntary Donor remove are the agent's own body. Accordingly, one might argue that in such cases the agents should be cut moral slack because, as it were, they

are only taking away with one hand a good given with the other.<sup>28</sup> This is not our argument.

Nor are we invoking the idea, often voiced in discussions of the Dif, that in cases like these, by failing to prevent deaths (albeit by causing them), the agents are merely “allowing nature to take its course.”<sup>29</sup>

Nor are we trying to argue that these cases are special because in them the final upshots of what the agent does are the same as if the agent had never existed.<sup>30</sup>

All of these ways of arguing try to justify our deontological intuitions by appeal to some sort of summation of the values of the *consequences* of an agent’s act or actions. Observe that following any or all of these lines of thought to their logical conclusion would require us to treat Interpose and Kick as on a par with Step Away and the rest. After all, in Interpose and Kick, the agent only removes a benefit—a barrier to the cart—that he himself has provided. He allows nature, in the form of the already careening cart, to take its course. And once the stone is kicked away, for all the difference the agent has made, he might never have been there at all.

From our point of view this way of thinking is deontology manqué. The Equivalence principle does not apply to Interpose and Kick. How the stone got there, what would have happened had the agent never existed...are beside the *deontological* point. The relevant deontological facts are those about how the agent acted. The agent was there and kicked the stone. Had he refrained from kicking, he would have allowed the child to live. We think he is a killer, no better than the agent in Push and we think that common sense would agree. Courts would convict him as readily as the agent of Drive Away.

But now what of Drive Away? Why does intuition treat this case as so different from Duck?

Consider how the agent acts. The agent moves his car out of the way. If it is an ordinary sort of car—and that is surely what our ready judgment assumes—then it must be that the agent makes the car move either by releasing the brake or pressing the throttle or both. In the case of ordinary cars only some such basic actions will get the

<sup>28</sup> Cf. McMahan, “Killing, Letting Die, and Withdrawing Aid.”

<sup>29</sup> Cf. Alan Donagan’s account in *The Theory of Morality* (Chicago: University Press, 1977); and Philippa Foot, “Killing and Letting Die,” in Bonnie Steinbock and Alastair Norcross, eds., *Killing and Letting Die* (New York: Fordham 1994, second edition), pp. 280–89.

<sup>30</sup> Shelly Kagan, *The Limits of Morality* (New York: Oxford, 1989), pp. 94–96.

things going. Suppose, in this case, the agent causes the car to move, and hence the child's death, by stepping on the throttle.

The agent's alternative was not to move the car. How could he have accomplished that? Well, by not stepping on the throttle. But not-stepping-on-the-throttle is no basic action and, however the agent performed it, no basic act of his would have been a cause of the car's motion or the child's death. So the driver's choice was between causing a death and allowing the child to live. The Equivalence Principle does not apply. The driver is as much a killer as the agent in Push, which is why the juries would convict.

Of course, this intuition is founded on the assumption that the car is an ordinary sort of car worked in an ordinary sort of way. What if it was an *extraordinary car*, a car rigged, say, so that only the driver's total immobility would keep it from moving? If that were the set up, the driver would cause the child's death by moving and his alternative would be to prevent the child from dying by total immobility. In that case, would not the Equivalence Principle apply?

It would. And, given these extraordinary details, we think no jury would convict. We can be confident at least that Jonathan Bennett would not convict. For, as now described, the case is—automotive elaborations aside—the Mobility case we considered early on. Bennett wanted to call that case an allowing. We disagreed, but now we can see why intuitions might run Bennett's way. The agent in Mobility causes the bad outcome *O*, but because his alternative is to prevent *O*, the Equivalence Principle applies. So Mobility is morally equivalent to an allowing. On the other hand, the Equivalence Principle *also* applies to the Immobility case, which may explain why Bennett's critics balked at counting it, as Bennett wanted to count it, as plain doing.

One could accept the Deontological Equivalence principle without accepting our account of the Dif. But to give the principle any purchase requires an account of the Dif which allows cases where the alternative to causing death may not be that of allowing life. Note that Bennett's account does not allow this. On Bennett's account, an agent makes an outcome happen if it is one of a minority of the ways the agent might have moved that have the outcome as an upshot. It follows that had the agent moved in one of the majority of ways which do not have that upshot, his action would be an allowing. Which is to say that on Bennett's account an agent's alternative to making an outcome occur is inevitably allowing it not to occur. Bennett's account is not unique in this. So far as we know, every account of the Dif other than our own entails that the alternative to killing must be allowing to live.

The Equivalence Principle morally equates actions given certain

alternatives; as we might expect, the equivalence holds for the alternatives as well.

*Deontological Equivalence 2:* Preventing a harm, when the alternative is causing that harm, is morally equivalent to preventing a harm when the alternative is allowing that harm.

Which, applied to death, gives us:

Preventing a person's death, when the alternative is causing that person's death, is morally equivalent to preventing a person's death when the alternative is allowing that person to die.

The second form of the principle deals with cases in which the agent refrains from causing a death by preventing it. For instance:

*No Step Away:* A cart is rolling down a hill. Agent is in its path at the top of the hill and if he remains where he is the cart will bump into him and stop. Agent realizes this and remains where he is.

Before appreciating the equivalence we might have thought that causing a death must always be as impermissible as it is in Push and that refraining from causing death must always be as obligatory as:

*No Push:* A cart stands at the top of a hill. If pushed it would roll down the hill and fatally injure a child. Agent does not push.

But the second Equivalence Principle tells us to treat the No Step Away agent as morally equivalent to the agent of No Stay Back:

*No Stay Back:* The cart is already rolling. Agent interposes a rock which stops it before it can hit the child.

The No Step Away and No Push agents both refrain from causing a child's death but they are morally different. The No Step Away agent is laudable; if the bump he received was sharp, maybe even heroic. Like the No Stay Back agent, he is a Good Samaritan. By not pushing, the agent of No Push has likewise refrained from causing a death. He has done the right thing but he is no Good Samaritan; he has done no more than any of us do when we refrain from manslaughter.

The mention of Samaritans will likely recall Judith Jarvis Thomson's 1971 article "A Defense of Abortion."<sup>31</sup> Of all our examples, Involuntary Donor bears a special seriousness since Thomson invented it as an analogy for abortion. Thomson conceded that by unplugging himself the agent kills, but she did not think it right to call it murder.

<sup>31</sup> Reprinted in William Parent, ed., *Rights, Restitution, and Risk: Essays in Moral Theory* (Cambridge: Harvard, 1986), pp. 1-19.

She tried to make her case in the language of rights. She argued that the victim has a right to life, but not a right to be saved and wanted to say that by unplugging, the involuntary donor was only failing to save, not killing, or anyway not killing unjustly. The donor had only failed to be a Good Samaritan.

We have not appealed to rights or justice or samaritanism but have arrived at a consonant result. The Equivalence Principle vindicates Thomson's core intuition. When a pregnant woman acts to bring about an abortion, it is clear that the alternative to this causing of death is a prolonging of life; her action is morally equivalent to allowing a death.

This is not to say that abortion or allowing-to-die is morally blameless. For all we have said, those who think that allowing someone to die is morally equivalent to killing may continue to say so. But they are not entitled to say that they can not see the Dif.

KADRI VIHVELIN

University of Southern California

TERRANCE TOMKOW

Dalhousie University