

2009

## Blame, Praise, and the Structure of Legal Rules

Lawrence M. Solan

Follow this and additional works at: <http://brooklynworks.brooklaw.edu/blr>

---

### Recommended Citation

Lawrence M. Solan, *Blame, Praise, and the Structure of Legal Rules*, 75 Brook. L. Rev. (2009).

Available at: <http://brooklynworks.brooklaw.edu/blr/vol75/iss2/6>

This Article is brought to you for free and open access by BrooklynWorks. It has been accepted for inclusion in Brooklyn Law Review by an authorized administrator of BrooklynWorks. For more information, please contact [matilda.garrido@brooklaw.edu](mailto:matilda.garrido@brooklaw.edu).

# Blame, Praise, and the Structure of Legal Rules

*Lawrence M. Solan*<sup>†</sup>  
*In Honor of Roger Shuy*<sup>\*</sup>

## INTRODUCTION

This essay discusses the implications of some important work in experimental philosophy to criminal law and tort law. In a series of extremely interesting and challenging studies, Joshua Knobe and others have demonstrated that we are more likely to attribute intent to those whose actions lead to bad results than to those whose actions lead to good results. In one version, a middle-management executive boasts that his new process will not only make money for the company, but also help the environment. His boss approves the new venture, but makes it clear that he doesn't care about helping the environment, he's doing it for the money. Sure enough, the new process does help the environment. Yet we do not give the boss credit for intentionally having made the environment cleaner. In contrast, when the mid-level manager laments that the new process will make money at the expense of damaging the environment, we attribute to the boss who approved the new venture an intent to have harmed the environment.<sup>1</sup> This asymmetry is potentially important. It suggests that we are somehow designed to lash out at wrongdoers, making us all

---

<sup>†</sup> Don Forchelli Professor of Law, Associate Dean for Academic Affairs, and Director, Center for the Study of Law, Language and Cognition, Brooklyn Law School. Thanks to Joshua Knobe for discussion of many of these issues, to Steven Dean, Ted Janger, Adam Kolber, Rebecca Kysar, Bertram Malle, Brett McDonnell, Tony Sebok, Chris Serkin, Aaron Twerski, and Dan Simon for helpful suggestions, and to the University of Minnesota School of Law and the University of Southern California Gould School of Law, where I presented earlier versions of this work. Thanks also to Steven Bentsianov for his most valuable research.

<sup>\*</sup> I dedicate this essay to the career of my friend and colleague, Roger Shuy, Professor Emeritus, Georgetown University, who is a pioneer in the study of language and law, and whose life represents a commitment to sheer decency, the value discussed in this essay.

<sup>1</sup> See Joshua Knobe, *Intentional Actions and Side-Effects in Ordinary Language*, 63 ANALYSIS 190 (2003).

moral actors. If the asymmetry is the result of our attributing blame too easily, then it suggests that prosecutors and plaintiffs' lawyers might have it easier than the definitions of crimes and torts would suggest. It has received so much attention that it has been given a name: "The Knobe [E]ffect."<sup>2</sup>

In this essay, I suggest an explanation for the asymmetry by looking more closely at the language of intent and causation. Intent is built into our speech without our having to point it out. We mention intent in everyday communication only when a person's state of mind deviates sufficiently from baseline expectations. If you tell someone you went out for dinner last night, you do not need to tell them that you did it on purpose. By looking at the circumstances in which we do make mention of intent in ordinary speech, we can begin to draw inferences about what our baseline expectations are. I argue here that the Knobe Effect results from the fact that our baseline is that people will act within social norms to accomplish socially-useful goals. This means, somewhat ironically, that our bias toward blame results from our tendency to expect the good. To the extent that social norms track moral judgment, the asymmetry between blame and praise suggests that moral judgment affects cognitive judgment as Knobe claims.

As for causation, I agree with Knobe that moral considerations pervade our judgments. But they do so, as Mark Alicke<sup>3</sup> and others<sup>4</sup> have argued, when there are competing causal stories consistent with the facts, and we need to choose among them. Moral judgment influences our choice among possible, legitimate causal accounts, but it is not sufficient to create a causal account from one that would not otherwise be available.

In an earlier essay, I argued that blaming arises naturally because its primitives—the recognition of a bad outcome, causation, and attribution of state of mind—are such basic cognitive functions that moral attribution is cognitively

---

<sup>2</sup> See generally Adam Feltz, *The Knobe Effect: A Brief Overview*, 28 J. MIND & BEHAV. 265 (2008) for a discussion of some of the literature citing this asymmetry.

<sup>3</sup> Mark D. Alicke, *Culpable Causation*, 63 J. PERSONALITY AND SOC. PSYCHOL. 368 (1992) [hereinafter Alicke, *Culpable Causation*]. For Alicke's reaction to Knobe's work, see Mark D. Alicke, *Blaming Badly*, J. COGNITION & CULTURE (forthcoming) (manuscript at 2-4, available at <http://www.unc.edu/~knobe/Alicke.pdf>) [hereinafter Alicke, *Blaming Badly*].

<sup>4</sup> See, e.g., Dan M. Kahan & Donald Braman, *The Self-Defensive Cognition of Self Defense*, 45 AM. CRIM. L. REV. 1, 17-19 (2008).

inexpensive.<sup>5</sup> Knobe pointed out in response that my theory had no way of accounting for the kind of asymmetry between blame and praise illustrated by his corporate executives.<sup>6</sup> That is, if blaming is readily available cognitively because it is triggered whenever a person with a bad state of mind causes a bad outcome, and judgments about causation, valence, and state of mind are made routinely for other reasons in any event, then praising should be just as easy since it is triggered by exactly the same considerations, the only difference being the valence of the outcome. Knobe was correct.<sup>7</sup> This essay offers an explanation for the asymmetry and explores some legal ramifications.

Should the law care about whether we blame more easily than we give credit and the reasons for any such asymmetry? Whether or not it *should* care, it *does* care. For the most part, the legal system is structured around baseline expectations of appropriate conduct consistent with social norms, with civil or criminal sanctions following from failure to meet these norms. The system expects people to behave properly, and it reacts negatively when they do not. There is far less reward for good conduct than there is punishment for bad. Furthermore, the legal system puts a great deal of weight on causation and intent in areas such as criminal law and torts. To the extent that our judgments of causation and intent are at odds with the moral intuitions of the citizenry, the law stands to lose some of its moral force.

The first part of this essay consists of my effort to explain the Knobe Effect. The second touches on ways in which the asymmetry it describes is relevant to our understanding of how the law is structured. In particular, it is no accident that legal systems are generally organized around a set of high expectations for conduct and sanctions for falling below those standards, and not as an incentive system in which the expectations are low, and rewards are given for rising above them. While economists regard these approaches as equivalent, they are not the same from a psychological perspective. Further, the negative branch of the Knobe Effect—that we

---

<sup>5</sup> Lawrence M. Solan, *Cognitive Foundations of the Impulse to Blame*, 68 BROOK. L. REV. 1003 (2003).

<sup>6</sup> Joshua Knobe, *Cognitive Processes Shaped by the Impulse to Blame*, 71 BROOK. L. REV. 929, 930 (2006).

<sup>7</sup> Lawrence M. Solan, *Where Does Blaming Come From?*, 71 BROOK. L. REV. 939, 939 (2006).

construe unwanted side-effects as intentionally caused when we knew they would occur—plays an important role in tort law. In fact, Section 8A of the *Restatement (Second) of Torts* specifically calls for known negative side effects of an intentional act to be treated themselves as intentional acts.<sup>8</sup> However, the boundaries of this doctrine are currently a matter of dispute. Experimental work by Knobe and others demonstrates that at least some version of this principle is consistent with everyday moral intuitions. However, the *Restatement* distinguishes more than Knobe does between moral judgment on the one hand, and the attribution of actual intent on the other. Based on empirical findings by Bertram Malle and others, I suggest that the *Restatement's* position is probably the better one. The third part of the essay is a brief conclusion.

I. INTENT, PRAISE, AND BLAME

A. *Intent in Ordinary Speech*

What does it mean to do something intentionally? Psychologist Bertram Malle and Joshua Knobe conducted numerous experiments on the folk psychology of intention.<sup>9</sup> Basically, prototypical intent involves an individual with the skill to carry out a task performing that task because he intends to do it, knowing that he is engaging in that task, and wanting to achieve the predictable result of the action. Thus, intent is most often acknowledged when an individual intends not only his actions, but also intends the foreseeable consequences of his actions. In other words, the folk psychology of intent includes both general intent and specific intent, to use the legal terminology.<sup>10</sup> This explains why law students find it strange when a tort or crime requires only general intent. They are not used to disaggregating their ordinary sense of what it means to intend an action.

---

<sup>8</sup> RESTATEMENT (SECOND) OF TORTS § 8A (1965).

<sup>9</sup> See Bertram F. Malle & Joshua Knobe, *The Folk Concept of Intentionality*, 33 J. EXPERIMENTAL SOC. PSYCHOL. 101 (1997); see also Alfred R. Mele & Paul K. Moser, *Intentional Action*, 28 NOÛS 39 (1994).

<sup>10</sup> The distinction is commonplace in legal analysis. For one recent statement, see *People v. Chance*, 189 P.3d 971, 981 (Cal. 2008) (Kennard, J., dissenting) (“A specific intent crime is an offense that requires the defendant to not only intend to do an act but to also intend to achieve a consequence, such as (in the case of assault) the intent ‘to commit a violent injury on the person of another’ whereas a general intent crime requires only that the defendant intend to do the act.” (citations omitted)).

This is not to say that all five elements must always be present for us to attribute intent. Subsequent experimentation suggests that some conditions are necessary, others prototypical.<sup>11</sup> For example, someone who plans to do a bad thing for a bad purpose and then finds himself having done this bad thing without being aware that he did it will be judged to have at least some intent. Nonetheless, when all of these elements are present, people are most likely to agree that an actor has acted intentionally.

One way of learning about how we understand intent is to examine how we talk about it. Let us look more closely at how we speak of intent in ordinary language. Consider an individual, Mario, who has a ticket for an early morning flight, say, 6:30 a.m. He wants to get to the airport an hour early, and knows from experience that it takes about fifteen minutes to get there by car. The day before, he calls a taxi service and asks for a cab to pick him up at his home at 5:15 a.m. and drive him to the airport. And that is just what happens. Upon arriving at the terminal, he pays the driver, including a tip, enters the airport, checks in at the counter, passes through security, buys a newspaper to read on the flight, has time for a quick cup of coffee at the airport's Starbucks, of which he takes advantage, and then boards the plane.

Every one of these events is intentional. But it would sound very strange—indeed, almost demented—to point that out. Thus, we do not say:

- (1) *Mario intentionally called a taxi service and purposely asked for a cab to pick him up at his home at 5:15 a.m.*
- (2) *Mario paid the driver on purpose and intentionally gave him a tip.*
- (3) *Mario got out of the car with intent and then intentionally entered the airport.*

*And so on.*

We do not say these things, because we assume that he did them on purpose without ever mentioning his state of mind. Linguistic pragmatics tells us what is so strange about (1)-(3). In everyday interaction, we assume that the speaker intends to advance the conversation in a cooperative way, and we draw

---

<sup>11</sup> See Bertram F. Malle, *Intentionality, Morality, and Their Relationship in Human Judgment*, 6 J. COGNITION & CULTURE 87 (2006).

inferences about the speaker's intent by evaluating what is said against this underlying assumption. That is why if Mario tells his wife Adela, "I have a meeting at 5:30" in response to her earlier remark, "I think we need milk," Mario will have understood Adela to have asked him to buy milk, and Adela will have understood Mario to have said that he will not be able to do so. This is true even though Adela never asks Mario to buy milk, and Mario never says that he cannot buy milk. It appears from this snippet that the store at which he would have bought the milk will close before he has a chance to get there and they both know that fact, even though neither of them has said anything about it. Grice's Cooperative Principle ("[m]ake your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged"),<sup>12</sup> and more specifically, his maxim of relation ("[b]e relevant"), capture this process.<sup>13</sup> Mario assumes that Adela mentioned the milk for a reason, and Adela assumes that Mario intended a relevant contribution to the conversation when he mentions the time of his meeting, and draws appropriate inferences about the milk.

The Cooperative Principle also explains what is wrong with saying that Mario intentionally did all the things he did to get to the airport. Intent is built into our models of these events, including both intent to perform the actions and intent to achieve the foreseeable results of those actions. Grice's maxim of quantity ("[d]o not make your contribution more informative than is required")<sup>14</sup> operates to inhibit us from expressing a state of mind that we already infer from the language and context. For this reason, any mention of Mario's intentional state of mind implies that he is acting with a level of intent in addition to the intent that we would ordinarily understand without further mention of his state of mind. Otherwise, there would be no reason to mention it at all.

This does not mean that intent must be part of the semantics of the verbs in these sentences, which include *call*, *ask*, *pay*, *give*, *get out of*, and *enter*. For example, I can unintentionally call a taxi by pushing the wrong speed-dial button on my telephone. Perhaps I thought I was calling a

---

<sup>12</sup> H. P. Grice, *Logic and Conversation*, in SYNTAX AND SEMANTICS 3: SPEECH ACTS 41, 45 (Peter Cole & Jerry L. Morgan eds., 1975).

<sup>13</sup> *Id.* at 46.

<sup>14</sup> *Id.* at 45-46.

friend to take me to the airport, but I inadvertently called the taxi service. In that case, I can say,

(4) *I inadvertently called a taxi service to take me to the airport.*

a. *Thank goodness that mistake didn't cause any trouble.*

b. *It cost me a fortune.*

In (4a) the error has led to a good outcome, in (4b) a bad one. In both versions, however, I have overridden the default assumption that I called the taxi service on purpose. Because I can express my non-prototypical intent by adding an adverb (“inadvertently” in this instance) without creating a contradictory statement, it appears that it is not the verbal semantics that is the source of the default assumption. Rather, the assumption comes from an internalized social psychology that tells us that people do things for a reason and intend the foreseeable consequences of what they do,<sup>15</sup> which is consistent with Malle and Knobe’s folk psychology of intent that includes both general and specific intent. This clearly holds for our hypothetical cab ride to the airport, where every mention of intent seems strangely redundant.

By the same token, we can highlight our intent when we speak of everyday events by stating it explicitly. But we do so principally when we mean something other than the ordinary baseline intent of accomplishing the intended goal, such as getting to the airport. For example, I may say (1) (repeated as (5)), but only under some specific and unusual circumstances:

(5) *Mario intentionally called a taxi service and purposely asked for a cab to pick him up at his home a 5:15 a.m.*

Say that out of loyalty and friendship, you offer to take Mario to the airport in your car at some ungodly hour. You tell me that fact, and I utter (5) in response. By *intentionally*, I would not mean the opposite of *accidentally*. Rather, I would mean that Mario has acted with more than the ordinary intent that goes into calling a cab, referring to his added goal of saving you the inconvenience of driving him to the airport. I could finish the sentence by adding, “precisely so that you would not even think of going so far out of your way.” In such cases, the very

---

<sup>15</sup> For a discussion of intent being built in to our internalized social psychology, see DANIEL C. DENNETT, *THE INTENTIONAL STANCE* 17 (1987) and RAY JACKENDOFF, *LANGUAGE, CONSCIOUSNESS AND CULTURE* 262-65 (2007).

expression of intent indicates that I am expressing volition beyond the normal intent of accomplishing the expressed goal.

*B. Praise, Blame, and Baselines of Intent*

Now let us return to our corporate executive who has approved a new project that will help the environment, but doesn't care that it will help the environment. The puzzle with which we began was to find an explanation for the fact that we do not give him credit for his assistance. At this point, we can see why it is that we do not give him credit for intentionally helping the environment. If we said nothing about his state of mind, we would assume that he intended the benevolent consequences of his action: He approved a profitable venture that will help improve the environment. When we utter (6) under ordinary circumstances,

(6) *The CEO approved a process that is expected to be both profitable and environmentally friendly.*

the baseline assumption is that there is an intention not only to approve the process but also to achieve the foreseeable results of approving the process.

Contrast this situation—the normal one—with Knobe's story of the corporate executive who goes out of his way to say that he doesn't care about helping the environment in deciding to go ahead with the new venture. That person has announced that his level of intent is less than the baseline, since he does not intend to achieve the positive foreseeable consequences of his action, and in fact, is callously indifferent to the good he is doing the world. It should not be the least bit surprising that people do not accept the statement that he intentionally helped the environment. When a person acts with more intent than the baseline, we make mention of it as relevant new information; when a person acts more or less at the baseline of intent, we acknowledge the intent if asked about it, although mentioning it overtly sounds redundant as described in the previous section; and when a person acts with less intent than the baseline, attributing intent to that individual is taken as a false statement.

We can similarly explain our reaction to the version in which the executive was callously indifferent toward hurting the environment. People hearing that version tend to attribute intent notwithstanding that the executive does not have hurting the environment as an affirmative goal. This again is a

function of deviation from baseline expectations. We assume that people act within social norms to achieve socially accepted goals. In most academic environments, the moderator of a presentation need not utter (7) to the audience when the speaker takes the floor:

(7) *I'm sure you will behave yourselves during the presentation and not make noises during the talk.*

We have no reason to believe that the members of the audience would do such a thing, although such a statement might be appropriate enough in a gathering, say, of teenagers with emotional problems where the defaults are different, or at certain political events. In those situations, it might even be appropriate to mention that a listener accidentally disrupted a talk by, say, knocking over his chair, if the baseline assumption is that disruptions will occur with intent.

The executive who does not care about hurting the environment has acted with more intent than the default level for a person causing harm, which is to do so inadvertently. Let us turn to examples of causation to see how the baselines for causing good and causing harm vary. Compare the following two sentences:

- (8) a. *Mario fixed the printer.*  
 b. *Mario broke the printer.*

Both *fix* and *break* are causative verbs: causation is built into their meanings. Only if Mario actually caused the printer to be fixed/broken would it be appropriate to utter the sentences in (8). Crucially, we assume that Mario acted intentionally in (8a), but not in (8b). In (8b), our baseline assumption is that the breakage was an accident. Our default is that we break things by accident and fix them on purpose. These assumptions may be overridden by circumstances. If Mario was in the middle of a violent tantrum, his breaking the printer may have been one of many intentional acts leading to a bad outcome. But we do not assume that he intentionally violated social norms without our providing some information that will override the default assumption. With bad outcomes, it is not redundant to speak of intent specifically. Thus, there is nothing strange about (9):

- (9) *Mario intentionally broke the printer.*

Since our baseline for bad outcomes is a state of mind less culpable than intent, attributing intent to Mario adds new

information, and thus does not violate the maxim of quantity by adding information that we already know. Similarly, some verbs are more loaded with intent than is *break*. If, for example, we were to substitute *smash* for *break* in (8b), we would be more inclined to believe that he did it on purpose in the first place. Expressing intent overtly when the verb is *smash* seems redundant, but not as redundant as when the outcomes are positive. Thus, it is easy to attribute intent to the executive who does not care about the harm he causes, and most people do just that.

Note that intent might not be the best characterization of the executive's actual state of mind. Were the choices expanded to those in (10), the results might have been different:

- (10) a. *The executive intentionally caused harm to the environment.*
- b. *While the executive did not intentionally cause harm to the environment, he is nonetheless blameworthy because he acted with an unacceptably callous attitude toward the foreseeable consequences of his decision.*
- c. *The executive did not intentionally cause harm to the environment.*

My sense of the story is most consistent with (10b). But if I had only options (a) and (c) to choose from, I might opt for (a) on the theory that it at least recognizes that the executive has acted with a state of mind that is blameworthy. In contrast, (c) might be consistent with my not assigning blame, which would fall beneath my moral standards.<sup>16</sup>

Recent studies by Guglielmo and Malle<sup>17</sup> report results consistent with this intuition. They presented subjects with Knobe's corporate executive story and gave them a choice of five characterizations of the CEO's state of mind. Only 10% chose "the CEO intentionally harmed the environment," whereas 70% chose "the CEO intentionally adopted a profit-raising program that he knew would harm the environment."<sup>18</sup>

---

<sup>16</sup> Adams and Steadman make a similar point in their pragmatic explanation of the Knobe Effect. See Fred Adams & Annie Steadman, *Intentional Action in Ordinary Language: Core Concept or Pragmatic Understanding*, 64 ANALYSIS 173 (2004); see also Alicke, *Blaming Badly*, *supra* note 3, manuscript at 9.

<sup>17</sup> Steve Guglielmo & Bertram F. Malle, Can Unintended Side Effects be Intentional?: Solving a Puzzle in People's Judgments of Intentionality and Morality (July 2009) (unpublished manuscript, [http://www.uoregon.edu/~sgugliel/pubs/Guglielmo&Malle\\_side\\_effects.pdf](http://www.uoregon.edu/~sgugliel/pubs/Guglielmo&Malle_side_effects.pdf)).

<sup>18</sup> *Id.* at 16 (internal quotation marks omitted).

Because there were five choices, chance performance was 20%, which means that subjects affirmatively rejected the description that the CEO harmed the environment intentionally.<sup>19</sup> Notably, notwithstanding that the subjects distinguished between intentional and knowing acts, they assigned very high levels of blame, suggesting that intentionally causing a negative side effect is tantamount to intent with respect to moral culpability, even if the two are analytically distinct states of mind.<sup>20</sup>

The philosopher Thomas Nadelhoffer conducted an experiment that provides some support for the baseline explanation.<sup>21</sup> After discussing Knobe's results, he presented forty college students with the following vignette:

Imagine that Steve and Jason are two friends who are competing against one another in an essay competition. Jason decides to help Steve edit his essay. Ellen, a mutual friend, says, "Don't you realize that if you help Steve, you will decrease your own chances of winning the competition?" Jason responds, "I know that helping Steve decreases my chances of winning, but I don't care at all about that. I just want to help my friend!" Sure enough, Steve wins the competition because of Jason's help.<sup>22</sup>

When asked, 55% of the students said that Jason had intentionally decreased his own chances of winning, a percentage far larger than those who received the positive version in Knobe's study, but far smaller than those who received Knobe's negative version.<sup>23</sup> These mixed results should not be surprising. I personally vacillate between understanding the story as one about the infliction of self-harm beyond the baseline (accidental, as in getting a paper cut), and as one about Jason's taking a risk that he knew might mature into harm, but hoped that it would not. In both understandings, the question is how Jason's decision matches up to what we would consider ordinary judgment. Hence the absence of consensus.

If I am right about what I have said thus far, praise and blame operate symmetrically after all. Only when we deviate from the baseline do we notice intent in either direction. What creates the asymmetry is not that we think differently about

---

<sup>19</sup> See *id.* at 14-15.

<sup>20</sup> See *id.* at 26.

<sup>21</sup> Thomas Nadelhoffer, *On Praise, Side Effects, and Folk Ascriptions of Intentionality*, 24 J. OF THEORETICAL & PHIL. PSYCHOL. 196 (2004).

<sup>22</sup> *Id.* at 209.

<sup>23</sup> *Id.*

praise and blame, but that the baseline of intent is not half way between them. Knobe is thus correct in his claim that judgments of state of mind are colored by judgments of whether the individual is acting within acceptable norms. In turn, the social norms that best trigger the asymmetry may well be ones that contain moral content. Doris, Knobe, and Woolfolk adopt a similar perspective, accurately stating that “responsibility attribution is *from the outset* deeply infused with normative considerations.”<sup>24</sup>

Significantly, while the Knobe Effect results from the application of pragmatics to the baseline intent, it cannot be explained by pragmatics alone.<sup>25</sup> The key to triggering the effect is the baseline. What pragmatic implicature provides is evidence of the effect’s application in everyday discourse. Where we locate the baseline is not a matter of pragmatics—it is a matter of social norms. Once we identify the baseline, however, then the rest of the facts fall out from ordinary pragmatic implicature. We do not specify the state of mind that would be assumed in ordinary experience because that would be redundant, and thus a violation of Grice’s maxim of quantity. When state of mind is mentioned, it is because the state of mind is other than that which would be expected, and is thus providing new information. Our corporate executive who approved the environmentally-friendly process acted with *less* than the ordinary amount of intent, so it would be especially infelicitous for us to credit him by using language that assumes that he acted with *at least* the ordinary amount of intent.

This analysis differs from other efforts to offer pragmatic explanations for the asymmetry between blame and praise. In particular, Adams and Steadman argue that the attribution of intent in the story with a negative outcome results from people’s disapproval of the conduct and expressing that disapproval through intent as a pragmatic matter.<sup>26</sup> They argue:

We suspect that what is going on in the minds of the folk is that they disapprove of the chairman’s indifference to the harm of the

---

<sup>24</sup> John M. Doris, Joshua Knobe & Robert L. Woolfolk, *Variantism About Responsibility*, 21 PHIL. PERSP. 183, 197 (2007).

<sup>25</sup> See Joshua Knobe & Ben Fraser, *Moral Judgment and Causal Judgment: Two Experiments*, in MORAL PSYCHOLOGY (Walter Sinnott-Armstrong, ed.) (forthcoming).

<sup>26</sup> See Adams & Steadman, *supra* note 16, at 178.

environment. They want to blame that indifference and they know that their blame is stronger and more effective at discouraging such acts, if the chairman is said to have done the action *intentionally*.<sup>27</sup>

While there may be some bootstrapping of this sort, this explanation does not appear to be adequate to explain the range of data. For example, Steven Sverdlik has found that when given the opportunity, people are perfectly willing to find conduct both blameworthy and unintentional at the same time.<sup>28</sup> In Sverdlik's studies, this happens when someone knows that he will cause discomfort, would rather not cause that discomfort, but does it anyway, accepting the side effect of his actions.<sup>29</sup> A person is considered blameworthy for mowing his lawn early in the morning knowing that it will wake the neighbors, recognizing that he did this without specific intent.<sup>30</sup> For that matter, we find the consequences of negligent action to be blameworthy, without ascribing any intent at all.

### C. *Causation and Outcome*

*Help* and *harm* are causative verbs. Causation is embedded in their meaning. You cannot harm someone without causing that person to be worse off, and you cannot help a person without causing that person to be better off. As John Darley and I point out in an earlier article, people do not always agree upon whether a person caused an outcome or whether that person did not cause the outcome, but set the conditions that enabled the outcome to happen.<sup>31</sup> In an experimental study, we found that the more culpable an individual, the more people tended to consider that person the cause of a bad outcome. For example, when an individual left his keys in the ignition of a car, and the car was stolen by a teenager who in turn injured a pedestrian, people were more likely to call the key-leaver a cause of the accident when he intentionally left his keys in the car than when he did so innocently.

---

<sup>27</sup> *Id.* at 178.

<sup>28</sup> Steven Sverdlik, *Intentionality and Moral Judgments in Commonsense Thought About Action*, 24 J. THEORETICAL & PHIL. PSYCHOL. 224, 234 (2004).

<sup>29</sup> *Id.*

<sup>30</sup> *Id.* at 226-30.

<sup>31</sup> Lawrence M. Solan & John M. Darley, *Causation, Contribution and Legal Liability: An Empirical Study*, 64 L. & CONTEMP. PROBS. 265 (2001).

Mark Alicke reached similar conclusions.<sup>32</sup> He presented people with a story about an accident which may or may not have been the fault of the protagonist.<sup>33</sup> In one version, the protagonist was on his way home to give a gift to his parents, in another, to involve himself with drugs.<sup>34</sup> People were far more likely to blame him for causing the accident when he was up to no good, even though the description of the accident remained constant.<sup>35</sup>

Knobe and Fraser have conducted studies with similar results.<sup>36</sup> Two people with equal opportunity to service an office machine fail to do so: one is responsible for maintaining the machine, the other is not. We say that the failure of the responsible person caused the machine to break when it does so because it has not been properly maintained.

What the studies have in common is that they begin with stories in which causation is equivocal. In ours, it is equivocal because we can conceptualize the situation as being either causation or enablement. In Alicke's, causation is equivocal because subjects were confronted with a set of facts susceptible to different causal explanations. And in Knobe and Fraser's it is equivocal because the cause, if there was one, was the failure to interrupt a process that was otherwise occurring. In all three cases, subjects' moral judgment contributed to whether they believed that the protagonist had caused harm. Thus, while people appear able to judge intent and blame independently,<sup>37</sup> an actor's state of mind does influence judgments about causation when causation is equivocal, and judgments of causation influence the extent to which that person is held responsible.

Returning to Knobe's story of the corporate executives, just as people construe intent asymmetrically, they may construe causation asymmetrically as well. That is, if subjects had been asked whether the boss hurt or helped the environment without any mention of his state of mind, we might expect to see the same effect, although perhaps somewhat weaker. That is because, just as the Knobe Effect requires that there be some range of possible states of mind for

---

<sup>32</sup> Alicke, *Culpable Causation*, *supra* note 3, at 369.

<sup>33</sup> *Id.*

<sup>34</sup> *Id.*

<sup>35</sup> *Id.* at 370.

<sup>36</sup> See Knobe & Fraser, *supra* note 25.

<sup>37</sup> See Sverdlik, *supra* note 28, at 234.

the actor, the examples used to illustrate the phenomena also contain some uncertainty about causation. Moral judgment, it thus seems, contributes to attributions of both state of mind and causation when language and circumstance leave open more than one possibility.

## II. IMPLICATIONS FOR THE LEGAL SYSTEM

### A. *Laws as High Expectation Signals*

There is no doubt that the psychology behind our attributions of intent, causation, and moral responsibility is of legal relevance. As Robinson and Darley have noted, the legal system becomes suspect in the eyes of citizens when it uses definitions that are at odds with their everyday moral judgments.<sup>38</sup> I will not pursue this issue in detail here, but it is worth noting that the law's distinction between general and specific intent is quite unnatural, and might be the subject of interesting experimental research.

The asymmetry between the assignment of intent in good and bad outcomes may help to explain significant aspects of the legal system's structure. In the eighteenth century, Blackstone noticed in his introduction to the Commentaries on the Laws of England that legal systems tend to organize themselves around expecting people to behave within the legal rules, and punish them for failing to do so:

[W]e find that those civil laws, which enforce and enjoin our duty, do seldom, if ever, propose any privilege or gift to such as obey that law; but do constantly come armed with a penalty denounced against transgressors, either expressly defining the nature and quantity of the punishment, or else leaving it to the discretion of the judges, and those who are entrusted with the care of putting the laws in execution.<sup>39</sup>

The account of praising and blaming discussed in this essay predicts that this should be so. Our baseline expectation in everyday life is compliance with social norms, with the attribution of blame for those who fail to do so. It should not be at all surprising that we organize our legal system around the way we think about praise and blame in ordinary life as reflected both in language use and in judgment about moral

---

<sup>38</sup> PAUL H. ROBINSON & JOHN M. DARLEY, *JUSTICE, LIABILITY AND BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW* (1995).

<sup>39</sup> 1 WILLIAM BLACKSTONE, *COMMENTARIES* \*56.

attribution, as the work of Knobe and others illustrates. In fact, that legal systems are structured around defining illegal conduct, tempered by excuses and justifications, may well be universal.

John Mikhail's contribution to this Symposium suggests that this approach is on the right track. Mikhail (with the help of research assistants), has examined homicide statutes in 163 jurisdictions from around the world.<sup>40</sup> All 163 jurisdictions have homicide statutes, and all 163 statutes have a mental state element. In addition, some justifications, such as self-defense, and some excuses, such as diminished mental capacity, appear universal, although others, including duress, necessity, and provocation, are not. Nonetheless, the structure of the law—a basic prohibition, tempered by excuses and justifications—does seem to be characteristic of every legal system studied.<sup>41</sup>

This is by no means the only way for a legal system to organize itself. Economists point out that the withholding of a benefit for those who fail to engage in desired conduct has the same economic effect as giving the benefit generally, and imposing a sanction on those who do not engage in the conduct.<sup>42</sup> The tax system is replete with “tax expenditures” that have this structure. The most notable in today's news is the mortgage interest deduction, which encourages people to incur debt to become homeowners. Various kinds of energy tax credits also create incentives for people to engage in desirable behavior. As Professor Wolfman noted more than twenty years ago, illustrating his point with an example that is still current today,

If a special exclusion from an oil company's income were recognized as the equivalent of a subsidy (because taxpayers in other businesses are denied a similar exclusion), the congressional debate might focus more sharply on just who benefits, on whether the subsidy is wise

---

<sup>40</sup> John Mikhail, *Is the Prohibition of Homicide Universal?: Evidence from Comparative Criminal Law*, 75 BROOK. L. REV. 497 (2009).

<sup>41</sup> *Id.*

<sup>42</sup> The point has also become part of the legal discourse. See, e.g., Christopher Serkin, *Existing Uses and The Limits of Land Use Regulations*, 84 N.Y.U. L. REV. (forthcoming 2009) (“Of course, as many have observed, it is hard to find a limiting principle that distinguishes between a regulation preventing a harm and a regulation conferring a benefit, but that such a line even needs to be drawn demonstrates that existing uses are normally protected.”).

policy and, if so, on whether Congress can better achieve its objective through direct expenditure or through the tax system.<sup>43</sup>

Yet people do not experience the imposition of a tax the same as the withholding of a benefit. In a series of very interesting studies, Edward McCaffery and Jonathan Baron have explored some of these reactions by asking people to react to various scenarios relating to taxation and its economic equivalents.<sup>44</sup> As these authors point out, a tax bonus for those who have children is financially equivalent to a tax penalty for those who do not have children.<sup>45</sup> In both cases, people with children will pay less than people without children. What changes are the defaults. In one instance, a society says that we assume that you do not have children, but if you do, then you can pay less tax than the default. In the other, a society says that we assume that you do have children, but if you don't, you will have to pay a surcharge above the normal tax amount. Not surprisingly, McCaffery and Baron found that people strongly prefer bonuses to surcharges. They are willing to accept a bonus for marriage or for children, but not a surcharge for being single or childless, even though they come to exactly the same amount in terms of dollars taxed.<sup>46</sup> The literature on behavioral law and economics discusses other ways in which people react asymmetrically to situations that are financially equivalent as the result of framing effects.<sup>47</sup>

These results are consistent with the explanation of the Knobe Effect presented in this essay. People are quick to blame individuals for conduct falling beneath the baseline of acceptable behavior. They are indeed willing to reward when conduct rises above the baseline, but the baseline is relatively high, making blame scenarios more prevalent. Not having children does not fall below any standard of acceptable social

---

<sup>43</sup> Bernard Wolfman, *Tax Expenditures: From Idea to Ideology*, 99 HARV. L. REV. 491, 493 (1985) (reviewing STANLEY S. SURREY & PAUL R. MCDANIEL, *TAX EXPENDITURES* (1985)).

<sup>44</sup> See Edward J. McCaffery & Jonathan Baron, *Thinking About Tax*, 12 PSYCHOL. PUB. POL'Y & L. 106 (2006).

<sup>45</sup> *Id.* at 114-15.

<sup>46</sup> *Id.*

<sup>47</sup> See, e.g., Daniel Kahneman & Amos Tversky, *Choices, Values, and Frames*, in *JUDGMENT AND DECISION MAKING: AN INTERDISCIPLINARY READER* 147, 150-53 (Terry Connolly, Hal R. Arkes, & Kenneth R. Hammond, eds., 2d ed. 2000). For discussion in legal contexts, see Russell B. Korobkin & Russell Ulen, *Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*, 88 CAL. L. REV. 1051, 1108 (2000) and Jeffrey J. Rachlinski, *Gains, Loses, and the Psychology of Litigation*, 70 S. CAL. L. REV. 113 (1996).

conduct in our society, however. Thus, people are reluctant to penalize others for not having children, although they are willing to reward those who do on the theory that raising children produces some social good beyond that which is required of all citizens.

We do not, however, generally structure our laws as incentives to avoid fines and prison sentences for obeying the law, even if they can be accurately characterized that way. Staying with the tax code, criminal laws do not say, "Whoever pays his taxes shall not be fined," nor do we give a bonus to people who pay their taxes, perhaps in the form of a periodic tax amnesty. Rather, we enact laws that punish people who break the law.

We could do otherwise, and do otherwise in other institutional settings. Staying close to home, many law schools have summer research stipend programs to encourage faculty to write productively. A law school that pays faculty members ten percent of their salary in exchange for writing an article during the summer months could, consistent with the framework of legal systems that I have described, raise the salaries of all faculty members, and then withhold 9.1% of the salaries of those faculty members who do not write during the summer. Although economically equivalent, the punitive system would be experienced as just that: a punitive system, and would be demoralizing. Thus, in some kinds of settings, we prefer incentives for good conduct to punishment for bad conduct.

Nonetheless, the principal generalization holds: legal systems are chiefly structured around punishment for the violation of legal rules. The fact that they could equivalently be stated in terms of incentives is irrelevant. This is a significant by-product of the Knobe Effect, which in turn, is a by-product of our setting baseline conduct on the side of high expectations of good conduct for good reason. Recognizing the source of this orientation might give policy makers more flexibility in deciding when to override it. The most natural way to think may not be the most productive way to think when it comes to such matters as the creation of an efficient and productive economic system.

#### *B. Knowledge of Side Effects as Intent in Tort Law*

The fact that we have a high baseline expectation explains in part why we prefer punishment for bad conduct to

reward for good conduct in our legal systems, but it says nothing about whether a legal system will reflect Knobe's finding with respect to those whose acts lead to bad side effects: that the legal system will attribute intent. It is entirely possible to structure a legal system around punishment, but not hold those who knowingly cause negative side effects responsible for having done so intentionally.

Significantly, the law generally acts in conformity with this aspect of Knobe's work as well. The law does not apply evenly, but in broad strokes it reflects the moral intuitions of Knobe's experimental subjects. This is an important point, because, as Robinson and Darley argue forcefully, respect for the criminal law is in large part contingent upon its reflecting the everyday intuitions of the citizens to whom it applies.<sup>48</sup> To the extent that the law treats lightly those who act intentionally, knowing—but not intending—the bad side effects of their acts, the law will be seen as encouraging immoral behavior. Consider Section 8A of the *Restatement (Second) of Torts*:

The word "intent" is used throughout the Restatement of this Subject to denote that the actor desires to cause consequences of his act, or that he believes that the consequences are substantially certain to result from it.<sup>49</sup>

Comment b then explains:

All consequences which the actor desires to bring about are intended, as the word is used in this Restatement. Intent is not, however, limited to consequences which are desired. If the actor knows that the consequences are certain, or substantially certain, to result from his act, and still goes ahead, he is treated by the law as if he had in fact desired to produce the result.<sup>50</sup>

The *Restatement* provides the following illustration:

1. A throws a bomb into B's office for the purpose of killing B. A knows that C, B's stenographer, is in the office. A has no desire to injure C, but knows that his act is substantially certain to do so. C is injured by the explosion. A is subject to liability to C for an intentional tort.<sup>51</sup>

---

<sup>48</sup> See ROBINSON & DARLEY, *supra* note 38, at 201-02.

<sup>49</sup> RESTATEMENT (SECOND) OF TORTS § 8A (1965) (emphasis added).

<sup>50</sup> *Id.* cmt. b.

<sup>51</sup> *Id.* illus. 1.

Significantly, the *Restatement* does not say that knowledge and intent are indistinguishable. Rather, it says that those who act with knowledge that their act will cause harm are held to the same level of responsibility as those who do act with intent. Thus, the *Restatement's* position is consistent with the results of Guglielmo and Malle reported above,<sup>52</sup> and with my own analysis, which suggests that subjects recognize the difference between knowledge and intent, but judge acts with known negative side effects harshly because the conduct falls so far below the baseline of societal expectations.<sup>53</sup>

To see the application of this principle in more detail,<sup>54</sup> consider *Vision Air Flight Service v. M/V National Pride*,<sup>55</sup> a case decided by the Ninth Circuit in 1998. The plaintiff had sold two airplane refueling trucks to a buyer for use in an airport in the Philippines. The defendant company was responsible for shipping the trucks. When it removed the first of the two trucks from the ship using a crane, it damaged the truck severely. The cable harness that was put around the refueling truck tightened up around the vehicle and literally crushed it. That was a negligent act, and under the terms of the contract, the shipper was responsible only for \$500 per truck. But then the shipper off-loaded the second refueling truck in exactly the same manner, and destroyed that truck too.<sup>56</sup> Under a doctrine of admiralty law called "deviation," the \$500 liability limit does not apply if a carrier intentionally destroys the goods it contracts to carry,<sup>57</sup> so the issue in the case was whether the destruction of the second truck should be considered an intentional act.

Citing the *Restatement*, the court held that a person acts intentionally if the "consequences [of his act] are substantially certain" to occur.<sup>58</sup> After holding that the destruction of the first truck cannot be considered intentional as a matter of law, the court concluded:

---

<sup>52</sup> Guglielmo & Malle, *supra* note 17, at 24-26.

<sup>53</sup> See *supra* Part I.

<sup>54</sup> The classic illustration used in Torts casebooks is *Garret v. Dailey*, 279 P.2d 1091, 1094 (Wash. 1955), where the question was whether a child who pulled a chair out from under an adult who was about to sit down had committed a battery even though the child meant no harm.

<sup>55</sup> 155 F.3d 1165 (9th Cir. 1998).

<sup>56</sup> *Id.* at 1167-68.

<sup>57</sup> *Id.* at 1175.

<sup>58</sup> *Id.* at 1176 (internal quotation marks omitted) (quoting RESTATEMENT (SECOND) OF TORTS § 8A (1965)).

The same, however, cannot be said with respect to the second refueler. The severe damage to the first refueler that resulted from off-loading was visible upon removal of the cable strapping as the refueler sat on the pier. Nevertheless, the stevedores proceeded to off-load the second refueler in precisely the same manner as the first. Because the uncontroverted evidence suggests the stevedores were aware of the damage their chosen method of off-loading inflicted on the first refueler, a rational trier of fact might very well conclude that they knew the second refueler would suffer the same fate with substantial certainty.<sup>59</sup>

The court's ruling is consistent with Section 8A of the *Restatement (Second)* and with the observations of Knobe and others that our moral intuitions tend to consider an undesirable side effect (i.e., the destruction of the truck) to be as blameworthy as an intentional act when the actor knows in advance that it will happen and acts anyway. Thus, at least in this instance, tort law and everyday moral judgment seem to be in harmony.

But the law is not uniform in its treatment of this issue. First, as Kobick and Knobe point out in their contribution to this Symposium, courts do not always impose liability for known negative side effects.<sup>60</sup> They discuss a 2009 U.S. Supreme Court decision in which the Shell Oil Company was held not to have acted intentionally when pesticides carried as cargo leaked from its tanker, as routinely happens when chemicals are offloaded.<sup>61</sup> The EPA spent some \$8 million to clean up the spill, and then sued Shell under CERCLA to recover the cleanup costs.<sup>62</sup> The statute would hold Shell liable if it had "arranged for [the] disposal . . . of hazardous substances."<sup>63</sup> This, according to Kobick and Knobe is a classic case of a company desiring only to engage in profitable activity, but acting with knowledge that it will cause environmental harm as a side effect. Shell did not act with the purpose of polluting the environment, it acted for the purpose of delivering chemicals that it had sold. Nonetheless, it knew that the spill would occur, and delivered the chemicals anyway.<sup>64</sup>

---

<sup>59</sup> *Id.*

<sup>60</sup> Julia Kobik & Joshua Knobe, *Interpreting Intent: How Research on Folk Judgments of Intentionality Can Inform Statutory Analysis*, 75 *BROOK. L. REV.* 409 (2009).

<sup>61</sup> *Burlington N. & Santa Fe Ry. v. United States* 129 S. Ct. 1870, 1883-84 (2009).

<sup>62</sup> *Id.* at 1876.

<sup>63</sup> 42 U.S.C. § 9607(a)(3) (2006).

<sup>64</sup> Kobik & Knobe, *supra* note 60.

The Supreme Court ruled 8-1 in favor of Shell,<sup>65</sup> a ruling seemingly inconsistent with the moral intuitions uncovered in one experiment after another. Why would there be such consensus on a position that appears to be at odds with everyday intuitions? Perhaps the word “arrange” suggests a desire to accomplish the consequences. Had the statute merely required reimbursement for an actor that “intentionally disposed of hazardous substances,” the result would have been different. But this does not diminish Kobick and Knobe’s point. Rather, it suggests that it might have been Congress that acted outside ordinary moral standards, and the Court merely followed suit. Whichever branch of government is responsible for the gap between moral judgment and the law, the gap remains, as Kobick and Knobe demonstrate.

At the same time, some courts have imposed liability for intentional acts when the actor knew that the side effect would occur at some point in time, but did not actually know that it would occur in this particular instance. This liability for knowledge that harm will come to some victim, but without knowledge that it will come to the particular individual who suffered, goes beyond the intuitions reflected in studies that Knobe has conducted, and, for that matter, is controversial among courts and commentators.<sup>66</sup>

Consider one of Knobe’s scenarios: a lieutenant orders a sergeant to send his troops up Thompson Hill. The sergeant responds: “But if I send my squad to the top of Thompson Hill, we’ll be moving the men directly into the enemy’s line of fire. Some of them will surely be killed!”<sup>67</sup> Some of the men are killed once the troops ascend the hill, and consistent with Knobe’s other studies, participants in this study say that the lieutenant acted with intent in causing the soldiers’ deaths.<sup>68</sup> Knobe did not investigate how subjects would react to the question of whether the lieutenant intentionally caused the death of any particular soldier who died on Thompson Hill. In his scenario, the prediction that *some* troops will die matches the fact that *some* troops did die. The lieutenant surely knew that each individual soldier was at risk of dying, and that at least some of the soldiers would actually die. But he did not know if some

---

<sup>65</sup> *Burlington N. & Santa Fe Ry.*, 129 S. Ct. 1870.

<sup>66</sup> For discussion of the different positions taken by the various courts, see RESTATEMENT (THIRD) OF TORTS § 1 cmt. a (Proposed Final Draft No. 1, 2005).

<sup>67</sup> Knobe, *supra* note 6, at 932 (internal quotation marks omitted).

<sup>68</sup> *Id.* at 933.

individual soldier, call him Bill, would be one of the casualties. Whether there would be consensus that the lieutenant intentionally caused Bill's death remains an open question.

To illustrate, consider *Parret v. Unicco Service Company*, a 2005 case decided by the Supreme Court of Oklahoma.<sup>69</sup> An employee (Parret) of a maintenance company (Unicco) was electrocuted while replacing emergency lights at a Bridgestone tire plant in Oklahoma City. Bridgestone was a client of Unicco. Working on a lighting system that was not electrically disabled was considered dangerous by both the maintenance company and Bridgestone. Because Parret was killed on the job and in the course of his employment, Parret's wife was limited to the benefits afforded under Oklahoma's workers' compensation system. However, the Oklahoma worker's compensation statute has an exception for those who are injured or killed by an intentional act. Unlike other workers, they are entitled to sue in court for tort damages. Parret's widow sued Unicco and Bridgestone, alleging that the two companies should be held to have intentionally caused her husband's death.<sup>70</sup>

The case went to the state supreme court, which ruled that an act should count as intentional for purposes of workers' compensation law when the employer knew with reasonable certainty that the injury (or death in this case) would occur. The opinion is written narrowly in this respect:

Thus, the employer must have acted, or have failed to act, with the knowledge that injury was substantially certain, not merely likely, to occur. The employer must have knowledge of more than "foreseeable risk," more than "high probability," and more than "substantial likelihood." Nothing short of the employer's knowledge of the "substantial certainty" of injury will remove the injured worker's claim from the exclusive remedy provision of the Workers' Compensation Act, thus allowing the worker to proceed in district court.<sup>71</sup>

Regardless of how the court stated the rule, however, the question in that case was not whether the employer and client wanted him dead—of course they didn't. Rather, it was whether their knowledge of the risk that someone would die sooner or later if workers were ordered to do such dangerous work amounted to knowledge, which in turn would imply

---

<sup>69</sup> 127 P.3d 572 (Okla. 2005).

<sup>70</sup> *Id.* at 574.

<sup>71</sup> *Id.* at 579.

intent. This additional step in reasoning goes beyond the moral intuitions that Knobe tested.

New Jersey has moved even further toward risk and away from actual knowledge as the basis for attributing intent to negative side effects. *Laidlaw v. Hariton Machinery Co.*, a 2002 case,<sup>72</sup> involved an employee who was severely injured on the job when his gloved hand got caught in a machine into which he was feeding material as part of a manufacturing project. The employer had disabled the safety guard that could have avoided the accident, although it periodically put the guard in place when it knew that inspectors would be coming to the plant. As of the date of the accident, no one had been injured by the machine, although there had been at least one close call, and the injured employee himself had expressed concern about the machines being unsafe when operated by new, inexperienced employees.<sup>73</sup>

The Supreme Court of New Jersey held that the employer should not have been granted summary judgment because a reasonable juror could have found that the employer knew with substantial certainty that a serious injury might ensue from using the machine without the safety device.<sup>74</sup>

The injurer's state of mind in these two cases is subtly different. In the Oklahoma case, the employer assigned a task to a series of people that sooner or later was bound to injure or kill someone. In the New Jersey case, the same employee performed the task repeatedly until he was hurt.

Professor Sebok,<sup>75</sup> commenting on the criticism of these extensions of intentional tort in an article by Professors Henderson and Twerski,<sup>76</sup> refers to the scenarios as the "iterated low risk act" and the "iterated low risk victim."<sup>77</sup> Both articles comment on the Discussion Draft of the *Restatement (Third) of Torts*, which adopts the "substantial certainty" extension of intent.<sup>78</sup> Henderson and Twerski are concerned not

---

<sup>72</sup> 790 A.2d 884 (N.J. 2002).

<sup>73</sup> *Id.* at 887-88.

<sup>74</sup> *Id.* at 897.

<sup>75</sup> See Anthony J. Sebok, *Purpose, Belief, and Recklessness: Pruning the Restatement (Third)'s Definitions of Intent*, 54 VAND. L. REV. 1165 (2001).

<sup>76</sup> See James A. Henderson, Jr. & Aaron D. Twerski, *Intent and Recklessness in Tort: The Practical Craft of Restating the Law*, 54 VAND. L. REV. 1133 (2001).

<sup>77</sup> Sebok, *supra* note 75, at 1170 (internal quotation marks omitted).

<sup>78</sup> The then-current draft discussed in these articles was RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL HARM § 1 (ALI Reporter Discussion Draft Fall 1999).

with the basic notion that a person who knows that his act will result in injury should be held to have caused that injury intentionally, but rather with the failure of the *Restatement's* definitions to preclude courts from expanding that principle to include the iterated low risk victim and iterated low risk act cases.<sup>79</sup> Professor Sebok would solve the problem by eliminating the principle altogether: an intentional tort is a tort committed with intent. Basic negligence law should be good enough, he argues, to take care of the other cases.<sup>80</sup> As an alternative, Professor Simons suggests that the kinds of cases that generate these problems, such as the maritime cases and workers' compensation cases discussed above, should be treated separately from basic tort law, again permitting the concept of intentional torts to be reduced to a simple one in which the tort is committed with the consequence intended as well.<sup>81</sup>

In both of these problematic scenarios, the line that has blurred is the line between knowledge that harm will come to the plaintiff here and now, and knowledge that harm will come sooner or later (whether to this plaintiff or to someone else) if the activity continues. In essence, these cases occur on the border between intentional and reckless conduct, a notoriously vague border to characterize.<sup>82</sup> Thus, it should not be entirely surprising that some judges regard as intentional torts harm caused by egregiously reckless conduct. The comments to the Discussion Draft of the *Restatement (Third)* recognize the controversial nature of these extensions, and attempt to limit their impact.<sup>83</sup>

---

<sup>79</sup> See Henderson & Twerski, *supra* note 76, at 1136-37.

<sup>80</sup> See Sebok, *supra* note 75, at 1174-76.

<sup>81</sup> See Kenneth W. Simons, *A Restatement (Third) of Intentional Torts?*, 48 ARIZ. L. REV. 1061, 1096 (2006).

<sup>82</sup> See Sebok, *supra* note 75, at 1174-75; Geoffrey Christopher Rapp, *The Wreckage of Recklessness*, 86 WASH. U. L. REV. 111, 133-40 (2008).

<sup>83</sup> "The applications of the substantial-certainty test should be limited to situations in which the defendant has knowledge to a substantial certainty that the conduct will bring about harm to a particular victim, or to someone within a small class of potential victims within a localized area. The test loses its persuasiveness when the identity of potential victims becomes vaguer, and when in a related way the time frame involving the actor's conduct expands and the causal sequence connecting conduct and harm becomes more complex." RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL HARM § 1 cmt. e (Tentative Draft No. 1, 2001).

### III. CONCLUSION

Attributions of intent are indeed infused with moral judgment, but indirectly so. It is not the case that we decide asymmetrically whether someone acted intentionally depending on the good or harm that was caused. Rather, we attribute intent when an actor's state of mind equals or exceeds the baseline norm for intentional activity. This baseline, in turn, is not situated half way between acting for the good and acting for nefarious purposes. Rather, it is skewed highly toward morally defensible and useful conduct. The result is that we are likely to ascribe intent to side effects that fall below the baseline more readily because unwanted positive effects are themselves below the baseline, a point also made by the philosopher Thomas Nadelhoffer.<sup>84</sup> The baseline itself tells us that positive side effects should not be unwanted.

This account has serious ramifications for legal systems. First, it suggests that legal systems are likely to expect the good and disapprove of the bad, rather than expecting the bad and rewarding the good. If people generally set their baselines high, there is no reason for legal systems to do otherwise. This prediction is borne out.

Secondly, the negative branch of the Knobe Effect—that people do indeed attribute intent to unintended negative side effects when the actor knew that they would occur—should also find its reflex in the law if the law does a good job reflecting our moral intuitions. This prediction is borne out as well, although somewhat unevenly. As Kobick and Knobe point out, courts and legislatures sometimes let actors off the hook for the unintended but known negative consequences of their action.<sup>85</sup> On the other hand, some courts have imposed liability on actors for side effects that they did not know would occur, but which they knew *might* occur if the activity continued. This is an extension of the experimental results on moral judgment, and remains controversial within the legal community.

Despite these issues, which are themselves ripe for additional inquiry, the law of torts indeed treats as intentional the known unintended negative consequences of an act. Whether people actually construe the known negative side effects of intentional acts as themselves intentional or whether

---

<sup>84</sup> Nadelhoffer, *supra* note 21.

<sup>85</sup> Kobick & Knobe, *supra* note 60.

they recognize the difference between the two but attribute the same degree of moral blame, the law appears to be in harmony with these intuitions. At least in its black letter statement then, the law tracks moral judgment quite well. This relationship between legal doctrine and experimental results suggests that the work being conducted in experimental philosophy and cognitive psychology can indeed make a contribution to mainstream legal discourse. Whether or not we insist that the law track the moral judgments of the citizenry, at the very least we should wish to know the extent to which it does, and to assess the cost of any gaps between the two against the benefits of whatever social or economic good legal principles at odds with moral judgments might bring.