The Butler Problem Revisited
THOMAS NADELHOFFER

On the surface, it seems plausible that the goodness or badness of an agent's actions should be completely irrelevant to the question of whether she performed them intentionally, but there is growing evidence that ascriptions of intentional actions are affected by moral considerations. Joshua Knobe, for instance, has recently published a series of groundbreaking papers (2003a; 2003b; forthcoming) in which he suggests that people's judgments concerning the intentionality of an action may sometimes depend on what they think about the action – morally speaking. One of the more interesting results of Knobe's psychological experiments is the discovery that people may have a lower threshold for judging that lucky (or unskilled) actions are intentional when these actions are *praiseworthy* or *blameworthy* than they do for judging that equally lucky (or unskilled) *morally neutral* actions are intentional.

In this paper I show that this discovery – when supplemented with some additional empirical data – gives us a way of shedding new light on a controversy that was sparked by Ronald Butler in 1977 when he posed the following problem to the readers of *Analysis*:

> If Brown in an ordinary game of dice hopes to throw a six and does so, we do not say that he threw the six intentionally. On the other hand if Brown puts one cartridge into a six-chambered revolver, spins the chamber as he aims it at Smith and pulls the trigger hoping to kill Smith, we would say if he succeeded that he had killed Smith intentionally. How can this be so, since the probability of the desired result is the same? (1977: 113)

The most obvious difference between Brown's rolling a six and his shooting Smith is that rolling a six is morally neutral whereas killing Smith is immoral. Does this purely moral difference explain why we are tempted to say that whereas Brown did not intentionally

roll a six, he did intentionally shoot Smith, even though his chances of success and his relevant control over the outcome are the same in both cases?

Some commentators have suggested that moral considerations *do* explain the asymmetry of our intuitions. E.J. Lowe, for instance, asks, 'Why, then, are we inclined to say in the dice-game merely that Brown threw the six not-intentionally…and in the revolver case (incorrectly I would claim) that Brown killed Smith intentionally? The answer lies in the distinctive moral features of the two cases' (1978: 118). On his view, even though the moral difference between the two cases does explain why we have conflicting intuitions, the facts of the two cases – e.g. Brown's respective probabilities of success – are 'formally quite parallel' (1978: 118).

Other commentators deny that evaluative considerations act expansively on ascriptions of intentionality in the way that Lowe suggests. Leo Katz, for instance, says that while it is tempting to say that the different moral qualities of the two acts explain our competing intuitions, this temptation is misguided (1987: 189). Butler expresses a similar opinion with the following pointed remark: 'Far too many entrants argued that aiming a six-chambered gun at somebody raises questions of moral responsibility, and that this accounts for the disparity. Nonsense!' (1977: 113) By his lights, insofar as Brown could have just as easily shot a bottle intentionally under the same circumstances, the moral difference between Brown's rolling a six and his killing Smith cannot explain our clashing judgments about whether Brown acted intentionally in the two scenarios.

As we have just seen, there are at least two general ways of responding to Butler's problem: First, we could simply insist that despite the fact that the probabilities in the two scenarios are the same, there are nevertheless other salient differences between Brown's

rolling a six and his killing Smith that justify our disparate intuitions. If we take this route, we would need to identify the relevant asymmetry and explain how it justifiably affects our ascriptions of intentionality (e.g. Davies 1981; Ross 1978). Second, we could argue that the cases really are parallel and that the asymmetry of our judgments about them is unjustified. If we take this route, we would not only need to respond to those who adopt the former line of reasoning, but we would also need to explain the asymmetry of our intuitions (e.g. Lowe 1978; 1980; Stiffler 1981; 1982).

Alfred Mele and Steven Sverdlik (1996) offer a thorough and helpful analysis of Butler's problem that incorporates both of these aforementioned lines of reasoning. On their view, Butler's two original scenarios are not entirely symmetrical. Indeed, there are several significant differences. First, we are asked to compare Brown's rolling a six with his killing Smith. But this is a faulty comparison from the outset (see, also, Kramer 1978). After all, whereas Brown's rolling a six is more closely analogous to his firing the gun, his killing Smith is more closely analogous to his winning the dice game. This difference is noteworthy because it could be that while people intuitively believe that Brown did not roll a six intentionally, they may nevertheless believe that he did win the game intentionally. Incidentally, if Butler had asked whether Brown won the game intentionally and whether he killed Smith intentionally – the two respective outcomes of the scenarios – our intuitions may not have been as at odds with one another.

In order to determine whether this difference has an effect on people's intuitions, I conducted an experiment. Subjects were 120 undergraduates – each of whom received one of the following three cases:

Case #1 (C1):

> Brown wants to kill Smith now. So, he takes out his six-shooter, places a single bullet in the chamber and spins the chamber. After spinning the chamber, Brown takes careful aim at Smith from a distance of ten feet, pulls the trigger, and shoots Smith directly in the heart in just the way Brown hoped he would. As a result, Smith dies.
> Question: Did Brown intentionally kill Smith?
>
> Case #2 (C2):
> Brown is playing a simple game of dice. The game requires that Brown roll a six to win. So, hoping to get a six, Brown throws a die onto the table. Unluckily for the other players, the die lands six-up and Brown wins the game.
> Question: Did Brown intentionally roll a six?
>
> Case #3 (C3):
> Same as Case #2, except the students were asked:
> Question: Did Brown intentionally win the game?

As was expected, nearly all of the subjects who received C1 judged that Brown *did* kill Smith intentionally (92.5%) and nearly all of the students who received C2 judged that Brown *did not* roll a six intentionally (90%). This difference was highly statistically significant, $\_^2$ (1, N = 40) = 54.5, $p < .001$. So, on the surface, it appears that Butler was right to be puzzled by our competing intuitions about the two cases. After all, the disparity between the results of C1 and C2 is undeniable. But once we look at the results of C3, we find that 62.5% of the subjects said that Brown *did* win the game intentionally. Thus, when we compare like with like – i.e. C1 with C3 rather than C1 with C2 – we do not find the same marked asymmetry – 92.5% yes vs. 62.5% yes. And while this difference was also statistically significant $\_^2$ (1, N = 40) = 10.3, $p < .002$ – there was less asymmetry between C1 and C3 than C1 and C2. This suggests that the difference between Brown's rolling a six – a *means* to an end – and his killing Smith – an *end* and not a means – at least *partly* explains why Butler's problem creates conflicting intuitions.

Another possible asymmetry is that the two original scenarios involve actions – *viz*., trying to roll a six in order to win a game and trying to fire a gun in order to kill

someone – that are themselves not analogous. In an effort to correct for this possibility, Mele and Sverdlik construct analogues for Butler's two original scenarios. On their view, if we once again compare like with like, we should be able to further minimize the asymmetry of our intuitions. So, Mele and Sverdlik develop their own shooting scenario involving Brown, but instead of trying to kill Smith, in this scenario he is simply trying to win a game by shooting a target. And just as he did in Butler's original shooting scenario, Brown places a single bullet in a six-shooter, spins the chamber, aims the gun, pulls the trigger, and hits the target in just the way he had hoped. The only difference is that whereas in Butler's case Brown shoots Smith, in Mele and Sverdlik's non-moral analogue he shoots a bull's-eye.

In order to see whether Mele and Sverdlik's non-moral shooting case helps further minimize the asymmetry of peoples' conflicting intuitions about Butler's problem, I conducted another simple experiment. This time the subjects were 40 undergraduates – each of whom received the following case:

> Case #4 (C4):
> Brown is playing a game with a six-shooter and a paper target. The game requires that he place a single bullet in the chamber and spin the chamber before firing. After spinning the chamber, Brown takes careful aim at his target's bull's-eye from a distance of ten feet, pulls the trigger, and hits the bull's-eye dead center in just the way Brown hoped he would. As a result, Brown wins the game.

The subjects were then asked whether Brown shot the bull's-eye intentionally – 75% said yes. When we compare this result with those from C1 – which showed that 92.5% said that Brown shot Smith intentionally – we find that evaluative considerations *do* have *some* affect on folk ascriptions of intentionality even in the two shooting scenarios. After all, the *only* difference between C1 and C4 is the moral status of the object that is being shot; yet 17.5% more of the subjects in C1 judged that Brown shot Smith intentionally

than subjects in C4 judged that Brown shot the bull's-eye intentionally. This difference is statistically significant, χ^2 (1, N = 40) = 4.50, $p < .034$.

I was also curious whether moral considerations might affect people's intuitions about whether Brown intentionally rolled a six. If so, this would not only show that moral considerations sometimes trump considerations of skill, but it would also suggest that philosophers who claim that skill is a necessary condition of intentional action are incorrect (e.g. Mele and Moser 1994). Thus I conducted another experiment. This time I used Mele and Sverdlik's immoral die case. Subjects were 80 undergraduates – each of whom received one the following cases:

> Case #5 (C5):
> Brown wants to kill Smith now. Smith is in another building. There is a bomb in that building and Brown can detonate it only by producing a six-dotted image on the lens of a camera that is focused on the top of a table in Brown's room and wired to the bomb. So, Brown takes out a normal, fair, six-sided die and tosses it onto the table, hoping that it will land six-up. Unluckily for Smith, the die lands six-up. By throwing the six, Brown detonates the bomb, thereby killing Smith.
> *Question:* Did Brown intentionally roll a six?
>
> Case #6 (C6):
> Same as Case #5, except the subjects were asked the following question:
> *Question:* Did Brown intentionally kill Smith?

Much to my surprise, *55% of the subjects in C5 judged that Brown intentionally rolled a six*. Moreover, 87.5% of the subjects in C6 judged that Brown intentionally killed Smith.

When we combine the data from all of these Butler problem cases, we get the following results:

C1: Immoral shooting case: Did Brown intentionally kill Smith? Yes: 92.5%
C2: Non-moral die case: Did Brown intentionally roll a six? Yes: 10%
C3: Non-moral die case: Did Brown intentionally win the game? Yes: 62.5%
C4: Non-moral shooting case: Did Brown intentionally hit the target? Yes: 75%
C5: Immoral die case: Did Brown intentionally roll a six? Yes: 55%
C6: Immoral die case: Did Brown intentionally kill Smith? Yes: 87.5%

Overall, these results strongly suggest that moral considerations do fuel the asymmetry of our intuitions about Butler's problem. Indeed, once we compare non-moral cases with their immoral analogues as Mele and Sverdlik suggest, we find that in the shooting cases, for example, subjects in C1 were more likely to say that Brown shot Smith intentionally than subjects in C4 were to say that Brown shot the bull's-eye intentionally, even though all of the other factors in the two cases were identical. As we saw earlier, this asymmetry is statistically significant. And we also discover that moral considerations had an even more startling effect on the die examples. First, whereas 87.5% of the subjects of C6 judged that Brown killed Smith intentionally, only 62.5% of the subjects in C3 judged that Brown won the game intentionally. This difference is also statistically significant, $\_^2$ (1, $N = 40$) = 6.6, $p < .01$. Second, whereas 55% of the subjects in C5 judged that Brown rolled a six intentionally in the process of killing Smith, only 10% of the subjects from C2 judged that Brown intentionally rolled a six in the process of winning a game of dice. This difference is highly statistically significant, $\_^2$ (1, $N = 40$) = 18.46, $p < .001$.

Clearly this last result is the most puzzling of all, for how could the immorality of the outcome of Brown's rolling a six affect *whether or not he rolled the six intentionally*? In both cases, he hoped to roll the six, he had a reason for rolling six, and he tried to roll a six – to the extent that throwing the die on the table constitutes the only action that he could have performed to make rolling a six possible. Moreover, in both cases he lacked any control over the die and he did not have any skill that increased his chances of making the die land six-up. Indeed, as Mele and Sverdlik point out, assuming that the die is a fair one, 'and that in neither scenario does Brown have special power over the dice, it is clear, in both cases, that Brown does not intentionally throw a six. In that respect, these

two cases are parallel. If there is a relevant asymmetry, it must be found elsewhere' (1996: 280). The question is: Where should we look? Given that the results of C2 and C5 *do* reveal a statistically significant asymmetry in folk intuitions concerning whether Brown rolled a six intentionally, and provided this asymmetry *cannot* be attributed to differences in skill, it appears that the only explanation is that moral considerations do act expansively on folk judgments of intentionality. But how do they do so?

In "Acting Intentionally: Probing Folk Notions," Mele develops an error theory that is intended to explain how moral considerations affect folk ascriptions of intentionality (2001; see, also, Mele & Sverdlik 1996). On his view, because the folk *erroneously* assume that if an agent is morally responsible for having A-ed, then the agent must have A-ed intentionally, they falsely ascribe intentionality to the actions of an agent whenever the agent tries to A and is morally blameworthy for A-ing. And because it is 'easy enough to show people that, upon consideration, they themselves would reject this assumption', their tendency to allow the moral features of actions to affect their ascriptions of intentionality is both mistaken and corrigible (Mele 2001: 41). Thus, Mele suggests that we could disabuse the folk of their mistaken reliance on moral considerations by reminding them of garden-variety cases of negligence and recklessness where the agent is still morally responsible. Once people are reminded of these types of cases, they will purportedly see that the assumption that all blameworthy actions must have been performed intentionally is 'false by their own lights rather than by the lights of an externally imposed theory' (Mele 2001: 41).

In order to insure that my subjects were not relying on the faulty theoretical assumption Mele identifies, I followed his advice and conducted another experiment.

This time, subjects were 40 undergraduates – each of whom got a two-page questionnaire. On the first page, subjects read the following case:

> Case #7 (C7):
> Bob got rip-roaring drunk at a party after work. When the party ended, he stumbled to his car and started driving home. He was very drunk at the time—so drunk that he eventually lost control of his car, swerved into oncoming traffic, and killed a family of five. (Mele 2001: 41)

The subjects were then asked whether Bob intentionally killed the family. They were also asked to rank his moral blameworthiness on a scale of 0 to 6 (0 being no blame, 6 being a lot of blame). And even though the mean rating for Bob's blame was M=5.1, 92% of the subjects still judged that Bob's behavior was *not intentional*. This shows that the subjects of C7 recognized that an agent *could* be blamed for her behavior *even if she did not perform the behavior intentionally*. On the second page of the survey, the subjects were presented with the immoral die case from C5. Yet, contrary to Mele's prediction, the subjects' responses were virtually the same as the earlier group's responses: 47.5% of the subjects still said that Brown intentionally rolled the six.[1]

These results show that we cannot simply explain away the fact that moral considerations affect folk ascriptions of intentionality by assuming that the folk are guided by mistaken theoretical assumptions. Indeed, when viewed in light of Knobe's earlier research, my Butler problem experiments offer further support for the claim that the folk concept of intentional action is sensitive to questions of right and wrong. Moreover, it appears that when it comes to people's judgments of intentionality, considerations of luck and skill can sometimes be trumped by moral considerations – which spells bad news for any analysis of intentional action that has skill as a *necessary*

[1]Joshua Knobe conducted a similar experiment to see whether the results of his Jake studies could be explained by Mele's error theory and he got very similar results (2003a).

condition. After all, in the immoral die case (C5), Brown not only *luckily* rolls a six, but he does so *without having any relevant skill whatsoever*.[2] Yet, because Brown rolled the six as a means of *killing Smith* – rather than merely as a means of *winning a game* – subjects were significantly more likely to judge that he rolled the six intentionally. These results not only speak against Butler's aforementioned insistence that questions of moral responsibility *cannot* account for the disparity of our intuitions about his original two scenarios, but they suggest that philosophers whose analyses of intentional action focus exclusively on non-evaluative considerations should broaden the scope of their investigation.[3]

The Florida State University
Tallahassee, FL 32306-1500, USA
tan02@fsu.edu

[2] In this respect, the results from C6—a case that involves an action that does not and cannot involve skill—go beyond the data from Knobe's Jake cases. After all, whereas Brown could not have been skillful with the die, Jake could have been skillful with the rifle!

[3] I would like to thank Joshua Knobe, Al Mele, and Virginia Tice for helpful comments on earlier drafts of this paper.

References:

Butler, R. 1978. Report on *Analysis* "problem" no. 6. *Analysis* 38: 113 – 14.

Davies, K. 1981. Killing people intentionally, by chance. *Analysis* 41: 156 – 59.

Katz, L. 1987. *Bad Acts and Guilty Minds*. Chicago: University of Chicago Press.

Kraemer, E. 1978. Intentional action, chance and control. *Analysis* 38: 116 – 17.

Knobe, J. 2003a. Intentional action in folk psychology: an experimental investigation. *Philosophical Psychology* Vol.16, No. 2: 309 – 24.

_____. 2003b. Intentional action and side-effects in ordinary language. *Analysis* 63: 190 – 94.

_____. (forthcoming). Intention, intentional action and moral considerations. *Analysis*.

Lowe, E.J. 1978: Neither intentional nor unintentional. *Analysis* 38: 117 – 18.

_____. 1980. Peacocke and Kraemer on Butler's problem. *Analysis* 40: 113 – 18.

Mele, A. 2001. Acting intentionally: probing folk notions. In *Intentions and Intentionality*, eds. B.F. Malle, L.J. Moses, and D.A. Baldwin, 27 – 43. Cambridge: MIT Press.

Mele, A. and P. Moser. 1994. Intentional action. *No_s* 28: 39 – 68.

Mele, A. and S. Sverdlik. 1996. Intention, intentional action, and moral Responsibility. *Philosophical Studies* 82: 265 – 87.

Ross, D. 1978. He loads the gun, not the dice. *Analysis* 38: 114 – 15.

Stiffler, E. 1981. Butler's problem again. *Analysis*: 216 – 18.

_____. 1982. Lowe on intentionality. *The Philosophical Quarterly* Vol. 32, No. 126: 70 – 73.