The Explanatory Structure of Moral Worth

Harjit Bhogal
University of Maryland

---

1   Rightness itself and Right-making Features

Some actions are morally right – there is a matching between the action performed and the facts about rightness. Only some of these actions have moral worth.[1]

Rightness Itself (RI)   Morally worthy actions are motivated by the fact that the action is right

Right-Making Features (RMF)   Morally worthy actions are motivated by the features of the action that make it right

It's not clear that RMF accounts properly exclude *coincidentally* right actions. This is the major motivation for RI accounts.

Bad Jean   Jean's friend missed her bus to work and frets over being late to an important meeting; coming late would be a great embarrassment to her. Wanting to spare her friend a major embarrassment, Jean gives her a ride. Let's assume that giving her friend the ride is the right thing to do in these circumstances and the fact that it spares her friend a major embarrassment makes it right.[2] However, Jean is so focused on saving her friend embarrassment that she would murder her friend's ex-boyfriend if that was the only way to save her friend embarrassment.[3]

Can we give a version of the RMF account that avoids coincidentality worries?

2   Coincidence

Roughly speaking a coincidence has two parts.[4]

1. A striking match between component events.[5]
2. Some 'disconnection' between those events
    ↪ Two natural approaches: *Modal* and *Explanatory*

Questions about moral worth fit this structure.

2.1   The Modal Approach

Perhaps a matching between events is non-coincidental if it could not easily have failed to hold.

↪ Problem: There can be modally robust coincidences

[1] I'm assuming (with much of the literature) that morally worthy actions are objectively morally right.

[2] A worry: This isn't the full set of right-makers. Sure, but this doesn't resolve the case. More on this later.

[3] Adapted from Sliwa (2016).

[4] See, e.g. Hart and Honoré (1985), Lando (2017), Bhogal (2020), Berry (2020)

[5] See Baras (2022) for a comprehensive recent discussion of strikingness. Giving an account of strikingness won't be necessary here.

- 31, 331, 3331, 33331, 333331, 3333331 and 33333331 are each prime but this is just a coincidence – 333333331 is not prime.[6]

- **Protons and Electrons** Protons are positively charged. Electrons are negatively charged. However, the absolute value of their charge is the same. Specifically, protons have a charge of $1.602176634 \times 10^{-19}$ coulombs, while electrons have a charge of $-1.602176634 \times 10^{-19}$ coulombs.[7]

More complicated modal accounts – for example, ones which add sensitivity-style conditions – fail too.[8]

2.1.1 The Pertinence Constraint

Distinguish moral worth and a broader evaluation of the agents character. What the agent would do in alternative situations is perhaps directly relevant to the character of the agent, but not the worth of their actual action.

More generally, only the motives that *actually* led to action are determine moral worth, not counterfactual ones.[9] (Though counterfactuals can be *evidentially* relevant.)

2.2 The Explanatory Approach

The explanatory approach looks more promising. In a coincidence the component events are explanatorily 'disconnected'.

But what is explanatory disconnection?

3 Formulating the RMF

The slogan of the RMF view is that an action has moral worth if it is motivated by features that makes the action right.[10]

Correlational RMF An agent ϕ-ing has moral worth if (i) the action has some feature, F, that makes it right and (ii) the content of their motivation to ϕ is that ϕ has feature F.

ϕ is right

S ϕs

F

S is motivated to ϕ by F
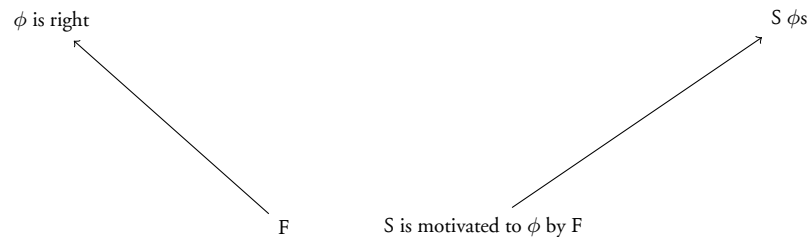
But there are counterexamples.

[6] Lange (2010)

[7] Bhogal (2023)

[8] See Bhogal (2023, section 3.2.2). Notice that sensitivity-style conditions on moral worth can seem too strong. Imagine S rightly prevents the deaths of thousands and that S would still prevent the those deaths even if the moral laws were different so that preventing deaths was bad. S's action can be worthy.

[9] See Isserow (2019), Markovits (2010, p. 210), Sliwa (2016, p. 399-400)

[10] Often in the moral worth literature 'motivation' is understood in a way that doesn't build in an explanatory connection between the action performed and the motivating fact. (Explanation is factive. So when authors describe cases where an agent is motivated by P but P is not true it seems that they are using this weaker sense of motivation.)
I'll follow this terminology.

Figure 1: The Correlational RMF view

Hiring Stephanie is a hiring manager and gives Yi-joon a job. The content of Stephanie's motivation is that Yi-joon is the most skilled programmer. Yi-joon is the most skilled programmer and that is a good reason to give them the job. However, Stephanie only believes that Yi-joon is most skilled because of incorrect racial stereotypes.

A natural fix is to demand that the agent's motivation is actually explained by the relevant right-making feature F. Call this the **Third-Factor RMF**.
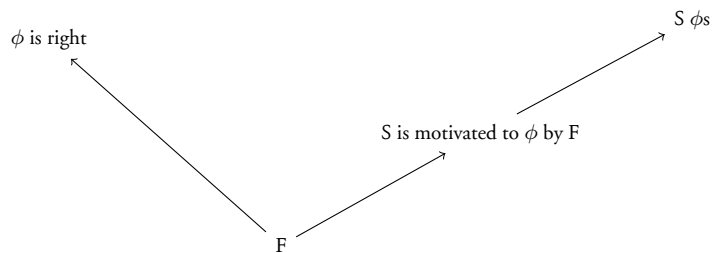

Figure 2: The Third-Factor RMF view

But **Bad Jean** is a counterexample.

4   STAPLING TOGETHER EXPLANATIONS

The problem is that these RMF views allow for cases that we explain why ϕ is right and why the agent did ϕ, but it still seems coincidental that the agent did the right thing. Sometimes it seems like we are merely *stapling together* explanations.

> One may explain each side of the coincidence in as much depth as one likes – going into wonderful normative depth about why family and friendship are valuable, and wonderful scientific depth about why we were selected to think this. But all this goes nowhere toward explaining the thing that really needs to be explained, namely the coincidence itself.[11]

[11] (Street, 2016, p. 31). See also Field (1996, section 5), Linnebo (2006, section 2) and Berry (2020, section 5b), among others.

Common Cause Pianos A boy is playing with two balls in the courtyard of an apartment complex. He throws both balls too high and they collide with each other in mid air and fly off in different directions. One of the balls hits a piano on one balcony and strikes a note. The other ball hits a different piano on another balcony and strikes a note. On each of the two pianos, the note struck is the high A.[12]

[12] Adapted from Lando (2017).

Clearly it's a coincidence that the same note was struck, even though the collision of the two balls is a common explainer of the two strikings.

To show some matching between A and B is non-coincidental we need an *explanation of the matching between A and B that is not merely an explanation of A and B together*. In the moral case we need an explanation of why the agent did the right thing that is not merely an explanation of why the agent did ϕ and why ϕ was right.

Unified Explanation RMF For an agent's doing ϕ to have moral worth (i) the agent must be non-instrumentally[13] motivated by right-making features of ϕ and (ii) there must be a *unified* explanation of why the agent did the right thing – that is, the best explanation of why the agent did the right thing cannot be just an explanation of why the agent did ϕ and why ϕ was right.

RI views can provide a unified explanation of the agent doing the right thing. If an agent has access to the moral facts and a desire to do the right thing that can explain why they act rightly without merely explaining why they did ϕ and why ϕ was right. This, I think, is much of the attraction of RI views.

But unified explanations of the agent acting rightly need not involve motivation by rightness itself. Some possibilities: 1. Moral insight 2. Good moral education 3. Secondary Motives.

5 Some Cases

5.1 Bad Jean

What is the explanation of why Jean did the right thing? In many versions of **Bad Jean** the best we can do is explain why Jean did ϕ and why ϕ is right. So Jean coincidentally does the right thing.

The fact that Jean would act wrongly in related situations is strong evidence that her actual action isn't appropriately connected to the moral facts. But it's not conclusive evidence – some versions of Bad Jean can act worthily, even if they would do terrible things in related situations.

5.2 Huck Finn

Huck is a white teenager living in south of the USA in the mid-19th century. He befriends an escaped slave, Jim. At a key point he is conflicted about whether to turn Jim in or to help him escape. He ends up helping Jim escape even though he thinks it is morally wrong since it amounts to stealing from Jim's 'rightful owner'.

5.3 Venom

Venom Jack, a surgeon, is hiking when he sees a stranger get bitten by a venomous snake and faint. He immediately makes an incision near the bite so that the venom will drain out. Making the incision is the right thing to do, and Jack's reason for doing it (that it will allow the venom to drain out) is part of what makes it right. But Jack doesn't have any particular concern for doing the right thing in this case, nor does he conceive of his reason as one that makes his action right. He is simply intrinsically interested in draining venom out of wounds.[14]

Many other RMF accounts face difficulties in both avoiding coincidence and not being so demanding that they rule out motivating cases like Huck Finn.

[13] This condition is to rule out cases like the politician who is motivated to help people in need, but only because being seen to do so will benefit his electoral chances.

[14] (Singh, 2020, p. 162)

References

Baras, D. 2022. *Calling for Explanation*. Oxford University Press.

Berry, S. E. 2020. Coincidence Avoidance and Formulating the Access Problem. *Canadian journal of philosophy*, 50(6):687–701.

Bhogal, H. 2020. Coincidences and the Grain of Explanation. *Philosophy and phenomenological research*, 100(3):677–694.

Bhogal, H. 2023. Explanationism versus Modalism in Debunking (and Theory Choice). *Mind*, 132(528):1005–1027.

Field, H. 1996. The a prioricity of logic. In *Proceedings of the Aristotelian Society*, vol. 96, pp. 359–379.

Hart, H. and Honoré, T. 1985. *Causation in the Law*. OUP Oxford.

Isserow, J. 2019. Moral Worth and Doing the Right Thing by Accident. *Australasian journal of philosophy*, 97(2):251–264.

Lando, T. 2017. Coincidence and Common Cause. *Noûs*, 51(1):132–151.

Lange, M. 2010. What Are Mathematical Coincidences (and Why Does It Matter)? *Mind*, 119(474):307–340.

Linnebo, . 2006. Epistemological challenges to mathematical Platonism. *Philosophical studies*, 129(3):545–574.

Markovits, J. 2010. Acting for the Right Reasons. *The Philosophical review*, 119(2):201–242.

Singh, K. 2020. Moral Worth, Credit, and Non-Accidentality. In Timmons, M. (ed), *Oxford Studies in Normative Ethics, Vol. 10*. Oxford University Press.

Sliwa, P. 2016. Moral worth and moral knowledge. *Philosophy and phenomenological research*, 93(2):393–418.

Street, S. 2016. Objectivity and truth: You'd better rethink it. *Oxford studies in metaethics*, 11(1):293–334.