Canadian Journal of Law & Jurisprudence XXXIX No.1 February 2026, 200-223

© The Author(s), 2026. Published by Cambridge University Press on behalf of Faculty of Law, Western University. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited. doi: 10.1017/cjlj.2026.10062

In Defence of Second-Order Reasons

Kenta Sekine
University College London, UK
Email: kentasekine1@gmail.com

Abstract

Some philosophers and legal theorists believe that certain fundamental normative phenomena cannot be adequately explained without appeal to second-order reasons—that is, reasons to act (or not to act) for certain reasons. Others are resistant to such an appeal. This resistance takes a robust form in the Credit Argument, which holds that, since we cannot act for a reason for a reason, the very idea of second-order reasons must be incoherent. In this paper, I do several things. Firstly, I clarify in just what sense the Credit Argument supposes it is not 'possible' to act for a reason for a reason. Secondly, I propose a novel typology of second-order reasons. Thirdly, using these insights, I demonstrate that the Credit Argument is weak at several points and should therefore be roundly rejected. This amounts to a partial defence of the appeal to second-order reasons in the explanation of fundamental normative phenomena.

Keywords: *acting for a reason; exclusionary reasons; moral worth; reasons for action; second-order reasons*

1. Introduction

In the theory of practical reasoning, it is generally accepted that considerations that have a normative bearing on action do so either positively or negatively. They speak either *for* or *against* that action; they are the *pros* and *cons* of doing it. Reasons for action must in this sense be either positive or negative reasons for action.[1]

It is also plausible that considerations can have a normative bearing on an action by bearing on doing it in certain ways. The fact that dinner is in the fridge might be a reason for going to the kitchen, but the fact that the baby is asleep in the kitchen might be a reason not to do so noisily. Situations like this are ordinary enough, and the description seems to be faithful to their normative structure. It seems right that the fact that the baby is asleep in the kitchen figures as a reason not to go there noisily, as opposed to a reason not to go there, full stop. After all, there may be other ways of going to the kitchen that will not wake the baby—quietly, for example.

Doing something quietly is, or at least can be, one way of doing it amongst many others. Plausibly, doing it for a certain reason is another of

1. Bare instances of 'reasons' refer to normative reasons unless otherwise specified.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

those ways.[2] If indeed it is, it follows that considerations can have a normative bearing on action by bearing on doing it for certain reasons. Considerations like this are what philosophers and legal theorists have in mind when they speak of *second-order reasons*. Like all reasons for action, they think, second-order reasons are either positive or negative: They are either reasons for acting for certain reasons, or reasons against acting for certain reasons.

The idea of second-order reasons was introduced to Anglophone philosophy and jurisprudence by Joseph Raz in his classic 1975 book *Practical Reason and Norms*.[3] One of the major projects of that book is to showcase the theoretical benefits of recognising second-order reasons by showing how a wide range of fundamental normative phenomena (e.g., decisions, norms, rules, obligations) cannot be adequately explained without appealing to negative second-order reasons—or as Raz calls them, 'exclusionary' reasons. Roughly put, the line of thought is as follows. What intuitively unifies the relevant phenomena is the special stringency they claim for themselves as reasons for action, a stringency they claim independently of their weight. We are required to do as they say, and not follow certain reasons to do otherwise, even if those countervailing reasons are in fact of greater weight. The special stringency thus seems to lie in how they take matters out of our hands—not by making us literally unable to pursue alternative options, but by normatively excluding doing so on their merits. It is this exclusionary force, a *sui generis* ability to defeat reasons to do otherwise irrespective of comparative weight, which is supposed not to be adequately explicable until we recognise second-order reasons. But once we do, the thought goes, we can say that the relevant phenomena defeat certain reasons to do otherwise in this *sui generis* way because they are negative second-order reasons—in this case, reasons not to act for those countervailing reasons.[4]

---

2. To do something 'for' a reason is to do it guided by that reason. For more on what it takes to be guided by reasons, see section 3. It is plausible that doing something for a certain reason is a way of doing it, because whenever someone does something for a certain reason, we can describe them faithfully as acting in the following way: responsively to that reason. Robert Audi accepts that acting for a reason is a way of acting, although admittedly he would deny that it is what he calls a "manner" of acting—for him a possible object of obligation—on grounds that we cannot choose the reasons for which we act. Robert Audi, "Acting rightly: Three dimensions of moral conduct" (2021) 34:1 Ratio 56 at 65-66. See section 3 for further discussion of this claim.

3. See especially Joseph Raz, "Postscript to the Second Edition: Rethinking exclusionary reasons" in *Practical Reason and Norms*, 2nd ed (Oxford University Press, 1990) 178.

4. Raz's explanation of the *sui generis* mode of 'exclusionary' defeat (*ibid* at 186-90) tends to get overlooked even by friends of exclusionary reasons. The idea, in brief, is to invoke an independently plausible principle that governs conflicts between reasons. Suppose there are two conflicting reasons, *A* and *B*, where *A* is weightier than *B*. Conforming with *A* completely precludes conforming with *B*, whereas conforming with *B* need not completely preclude conforming with *A* (e.g., because some other action constitutes partial conformity with *A*). In that case, we can conform with both reasons to some degree so long as we conform with *B*. Plausibly, there is a principle governing such conflicts that at least sometimes directs us to do whatever allows for the greatest measure of conformity with the applicable reasons, which in this case rules out conformity with the weightier reason *A*. Raz's thought is that the conflict between exclusionary and excluded reasons is like this. Since conforming with exclusionary reasons only requires that we not act for the excluded reasons, we can conform with both exclusionary and excluded reasons just so long as we conform with the excluded reasons without being guided by them. Thus the principle may apply, and if it does, then it requires us to do whatever

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Many have agreed with Raz on the need for second-order reasons to explain these normative phenomena (and others like them), and have allowed them to play a role in their various theories.[5] Others have been more resistant to second-order reasons, seeking to challenge their legitimacy.[6] The challenges fall into two broad groups.[7] *Theoretical* challenges say that the phenomena that allegedly cannot be explained without second-order reasons are either not in fact explained by them adequately, or amenable to explanation using less contested concepts.[8] But to mount a theoretical challenge is already in a sense a concessive move, since it grants that second-order reasons are coherent in the first place. *Conceptual* challenges target precisely this assumption.[9] There are indefinitely many ways to do this, in principle, but the dominant strategy has been to try to put conceptual pressure on the idea that we can act for a reason for a reason.[10] We explore the dominant strategy in more detail in what follows.

In this article I focus solely on conceptual challenges to second-order reasons. I pay special attention to the robust conceptual challenge advanced by Daniel

will allow for conformity with both reasons. That is, it rules out acting for the excluded reason, which it can do, we have seen, even if it is the weightier reason.

5. See e.g. John Gardner & Timothy Macklem, "Reasons" in Jules L Coleman, Kenneth E Himma & Scott J Shapiro, eds, *The Oxford Handbook of Jurisprudence and Philosophy of Law* (Oxford University Press, 2004) 440; Raimo Tuomela, "Group Reasons" (2012) 22:1 Philosophical Issues 402; Anna Stilz, "Occupancy Rights and the Wrong of Removal" (2013) 41:4 Philosophy & Public Affairs 324; Cass R Sunstein, "Beyond Cheneyism and Snowdenism" (2016) 83:1 U Chicago L Rev 271.
6. After an initial wave of challenges, followed by responses from Raz, there have been one or two subsequent waves that have gone largely without comment—from Raz or anyone else. Of the initial wave, see especially Michael S Moore, "Authority, Law, and Razian Reasons" (1989) 62:3&4 S Cal L Rev 827; Stephen R Perry, "Second-Order Reasons, Uncertainty and Legal Theory" (1989) 62:3&4 S Cal L Rev 913; Donald Regan, "Authority and Value: Reflections on Raz's *Morality of Freedom*" (1989) 62:3&4 S Cal L Rev 995; Larry Alexander, "Law and Exclusionary Reasons" (1990) 18:1 Philosophical Topics 5; Heidi Hurd, "Challenging Authority" (1991) 100:6 Yale LJ 1611; Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (Oxford University Press, 1991). For subsequent notable efforts, see Emran Mian, "The Curious Case of Exclusionary Reasons" (2002) 15:1 Can JL & Jur 99; Christian Piller, "Kinds of Practical Reasons: Attitude-Related Reasons and Exclusionary Reasons" in Sofia Miguens, João Alberto Pinto & Carlos E Mauro, eds, *Analyses: Proceedings of the 2nd National Meeting for Analytic Philosophy* (University of Porto Press, 2006) 98; Stephen Darwall, "Authority and Reasons: Exclusionary and Second-Personal" (2010) 120:2 Ethics 257; Daniel Whiting, "Against Second-Order Reasons" (2017) 51:2 Noûs 398; Noam Gur, *Legal Directives and Practical Reasons* (Oxford University Press, 2018). Points raised by Moore, Perry, and Regan are addressed directly in a response piece: see Joseph Raz, "Facing Up: A Reply" (1989) 62:3&4 S Cal L Rev 1153 [Raz, "Facing Up"]. Those raised by Darwall are addressed in Joseph Raz, "On Respect, Authority, and Neutrality: A Response" (2010) 120:2 Ethics 279. See also Raz, *supra* note 3; Joseph Raz, "Exclusionary Reasons" (last modified 14 October 2021), online: *SSRN* papers.ssrn.com/sol3/papers.cfm?abstract_id=3933033.
7. See Raz, "Exclusionary Reasons", *supra* note 6 at 10-11.
8. For a recent conspectus of both negative and positive theoretical challenges, as well as a novel positive theoretical challenge, see Gur, *supra* note 6.
9. Note that Raz himself raises a conceptual challenge to exclusionary reasons via a formal argument he calls "the basic objection." Raz, "Exclusionary Reasons", *supra* note 6 at 8ff.
10. We first find this dominant strategy in Moore, *supra* note 6 at 875-83. I believe that Raz's reply is enough to disarm Moore's challenge, so I make no attempt to relitigate that particular dispute here: see Raz, "Facing Up", *supra* note 6 at 1174-78.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Whiting.[11] In section 2, I explain how Whiting's challenge instantiates the dominant strategy and present his Credit Argument in the wider context of that strategy. I go on in sections 3-5 to argue that the Credit Argument has shaky premises, and that however charitably they are interpreted, the argument comes out either unsound or with a dramatically weakened conclusion. I conclude in section 6 that Whiting's challenge is unsuccessful.

2. Understanding Whiting's Challenge

The dominant strategy aims to put conceptual pressure on the idea that we can act for a reason for a reason.[12] Instances of the strategy can be thought of as having basically two components: first, a *principle* that licenses the inference from the existence of second-order reasons to our ability to act for a reason for a reason; and second, an *argument* that establishes the incoherence of our possessing such an ability. In this section, I bring out this bipartite structure in Whiting's challenge. Subsection 2.1 spells out the principle, subsection 2.2 the argument.

---

11. See Whiting, *supra* note 6 at 405-08. Note that this challenge to second-order reasons is not the only one Whiting raises in his wide-ranging paper. He offers intuitive doubts about the psychology of acting for second-order reasons too, although he concedes that these are "not universally shared" (*ibid* at 404). See section 3 of this paper for discussion of psychology. Whiting also objects to second-order reasons on the basis that they are of the "wrong kind" (*ibid* at 408ff). I find this objection doubly problematic. Firstly, I doubt that second-order reasons can be of the "wrong kind," as I agree with Ulrike Heuer that reasons for action, unlike reasons for attitudes, do not obviously admit of the philosophically interesting distinction between 'right' and 'wrong kinds' of reasons: see Ulrike Heuer, "Beyond Wrong Reasons: The Buck-Passing Account of Value" in Michael Brady, ed, *New Waves in Metaethics* (Palgrave Macmillan, 2011) 166 at 170-01. Secondly, and recognising this perhaps, Whiting proceeds on the very controversial assumption that "[i]n general, a reason to φ is a reason to decide or intend to φ"—an assumption I do not accept. Whiting, *supra* note 6 at 403[footnote omitted]. Hence, I disregard this objection in this article, although for a recent discussion that takes seriously the claims that reasons for action are reasons to intend, and thus that reasons for action admit of right and wrong kinds, see Ezequiel Monti, "(Really) defending exclusionary reasons" (2024) 15:1 Jurisprudence 48. The conceptual challenge I focus on is, I believe, more serious. Adams takes its success at face value as a partial basis for rejecting, or at least substantially revising, Raz's view of second-order—and thus exclusionary—reasons: see NP Adams, "In Defence of Exclusionary Reasons" (2021) 178:1 Philosophical Studies 235. By contrast, Keeling rebuts the challenge, although her view seems to foreclose guidance by reasons (*a fortiori* second-order reasons) *de re*, which I believe unnecessarily precludes the creditworthiness of agents like Huck Finn: see Sophie Keeling, "Responding to second-order reasons" (2024) 109:3 Philosophy & Phenomenological Research 799. See section 3 for how the view I propose avoids this implication.

12. Here, and in what follows, I take the referents of words like 'acts' and variables like 'φ' to include omissions. I also assume that reasons against actions can be understood as reasons for the corresponding omissions. This is common in the literature, although *cf* John Gardner & Heike Jung, "Making Sense of Mens Rea: Antony Duff's Account" (1991) 11:4 Oxford J Leg Stud 559; Justin Snedegar, "Reasons for and reasons against" (2018) 175:3 Philosophical Studies 725. Expressions like 'reasons to act for reasons' and their cognates are thus used to say things about second-order reasons in general (i.e., both positive and negative) without littering the text with parenthetical qualifiers.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

2.1. Principle (LR)

The principle component may differ in content for different instances of the dominant strategy, but its general form—its gist—can be expressed as the following schematic claim about reasons for action:

(R) If *p* is a reason for *S* to φ, then it is possible for *S* to φ for *p*.

Something like schema (R) looks highly plausible. It captures the widespread intuition that it is in the nature of reasons for action that they can guide our action.[13] If schema (R) is true, and we grant the existence of second-order reasons, it entails that we can act for a reason for a reason. To see this, consider that, in general, a second-order reason takes the form of a reason for *S* to ψ-for-*q*. Hence, if we make *p* in schema (R) a second-order reason by plugging ψ-for-*q* into the slot for φ, then it will follow, if (R) is true, that it is possible for *S* to ψ-for-*q* for *p*, i.e., that *S* can act for a reason, *q*, for a reason, *p*.

We can specify schema (R) in different ways depending on how we read 'possibility' in its consequent, and many specifications state sensible-looking principles governing reasons for action. But proponents of the dominant strategy are constrained in their choice of principle by the bipartite structure of their strategy. If the preferred principle gives too trivial a reading to 'possibility', then even if it is true, it will be tricky to make the argument that it is not possible, in the same sense, to act for a reason for a reason. For example, a metaphysical reading yields a true specification of schema (R), but it is not obvious how one would go about arguing that acting for second-order reasons is metaphysically impossible in the way, say, that giving birth to one's own father seems to be. Physical readings come in degrees of triviality, although note that on any such reading, the corresponding impossibility argument will presumably not fall in the philosopher's remit. At one extreme, we could mean 'physical possibility' in the sense of 'not excluded in principle by fundamental physical laws'. Like metaphysical readings, such physical readings seem too trivial to allow for a plausible impossibility argument: After all, lots of fanciful possibilities might not be so excluded (e.g., flying unicorns), so it is not obvious why it should be any different for acting for second-order reasons. At the other extreme, by 'physical possibility' we could mean 'causal possibility for beings physically constituted as we are', in which case we are referring to the psychological possibility of acting for second-order reasons. Such psychological readings are, conversely, so committed as to leave the resultant principle at risk of losing its veneer of truth. In that case, it becomes too easy to contest whether the associated impossibility argument speaks to the coherence of second-order reasons at all, even if the argument itself is compelling.[14]

---

13. See e.g. Christine M Korsgaard, "The Normativity of Instrumental Reason" in *The Constitution of Agency: Essays on Practical Reason and Moral Psychology* (Oxford University Press, 2008) 27 at 31; Joseph Raz, *From Normativity to Responsibility* (Oxford University Press, 2011) at 28.

14. This is in effect the trouble with Moore's version of the dominant strategy. Moore endorses a Hartian specification of schema (R), where 'possibility' is read in terms of having both the

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Whiting's preferred specification of schema (R) sits in a sweet spot between these extremes: committed enough to enable the construction of an argument, but not so committed as to be controversial. He calls it the "*motivational constraint*" (MC) on reasons, which at first glance looks the same as schema (R): "If that p is a reason for you to φ, it is possible for you to φ for the reason that p."[15] But Whiting goes on to specify a reading for 'possibility' in (MC). First of all, he says that "[t]o endorse (MC) is not to endorse the controversial 'internal' view of reasons."[16] He then adds that "[t]o endorse (MC) is not to endorse the controversial view that there is an 'epistemic filter' on reasons."[17] Putting these qualifications together, we can say that for it to be relevantly possible for *S* to φ for *p*, *S* need not know that *p*, nor need *S*'s subjective motivational set be such that they would or could conclude that they ought to φ by a sound deliberative route, were they to know that *p*.[18]

'Possibility' in Whiting's (MC) is thus very weak. To see this, suppose that *S* is completely ignorant of the facts, that they are entirely devoid of motivation, and that their doing φ would not anyway be conducive to realising whatever value there may be in satisfying *p*. Even in that case, it could still be relevantly possible for *S* to φ for *p*. Given the existence of a person *S* and the fact that *p*, there is no way the world could contingently be—i.e., as regards *S*, the wider facts, or their relations—that would rule out that it is possible for *S* to φ for *p* in the sense at hand. To highlight this point, we will call this a purely 'logical' sense. 'Possibility' in Whiting's (MC) means *logical possibility* in just this sense. In other words, his preferred specification of schema (R) is more perspicuously expressed as follows:

(LR) If *p* is a reason for *S* to φ, then it is *logically* possible for *S* to φ for *p*.[19]

Principle (LR) may be resisted due to the obscurity of 'logical' possibility. Admittedly, it is hard to get a grip on the sense in which it is 'possible' for *S* to φ for *p*, such that the possibility obtains independently of how things could

capacity and a fair opportunity to act for a reason: see Moore, *supra* note 6 at 875-76. What this means, for Moore, is that the existence of second-order reasons would entail that those to whom they apply are able to rationally choose the reasons for which they act, which he argues is inconsistent with the concept of causation. When Raz replies that second-order reasons entail only rational control over the reasons for which we act, as opposed to full-blown rational choice, he is contesting the Hartian specification of schema (R), or at least Moore's interpretation of it: see Raz, "Facing Up", *supra* note 6. This is why Raz can agree with the conclusion of Moore's argument, as he does when he writes, "we cannot choose but we can control" (*ibid* at 1177). See section 3 for more on the psychological possibility of acting for second-order reasons, and on the bearing of 'direct' and 'indirect' forms of control on such possibility.

15. Whiting, *supra* note 6 at 403.

16. *Ibid* [footnote omitted].

17. *Ibid* [footnote omitted].

18. This is not to say that for it to be relevantly possible for *S* to φ for *p*, *no one*—not even an ideal agent—need have the relevant knowledge or subjective motivational set, since this would make (MC) highly implausible. It is only to say that the one to whom the reason *p* applies, viz., *S*, need not have it.

19. This principle is considered by Gardner and Macklem in a perceptive paper to which the following discussion is greatly indebted: see Gardner & Macklem, *supra* note 5 at 455-57.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

contingently be with the relevant bits of the world. To do so, it will be helpful to reflect on the different ways it can be self-defeating for *S* to try to φ for *p*. In particular, I claim, it is logically possible for *S* to φ for *p* if and only if it is not necessarily self-defeating for *S* to try to φ for *p*. We can illustrate 'necessary self-defeat' with the following example.

Ben plays saxophone in an improv group. He has high ambitions for audience experience, always looking to deliver a powerful moment to remember. That it would lend spontaneity to the performance may seem to be a reason for him not to try, in performing, to construct the powerful moment he is after. There really would be value in the performance being spontaneous: The musical result would be less derivative, more ecstatic, and ultimately more powerful. But if Ben plays with one eye on this enhanced musical potency, always acting on this basis, then what he will be doing is precisely: trying to construct a powerful moment. Aspects of the value of musical performance have put Ben in a bind. It is not possible for Ben not to try to construct a powerful musical moment if he acts on the basis that it would lend spontaneity to his performance. For if he acts on the basis of this consideration, he *necessarily* fails to do as it would have him do—namely, not to try to construct a powerful moment. There is no way things could contingently be with the relevant bits of the world that would make it possible for Ben to do as the consideration would have him do on its basis. In this sense, the bind is a logical one. It is logically self-undermining for Ben to try to φ for *p*, and thus logically impossible for him to φ for *p*. This, together with (LR), is why the fact that it would lend spontaneity to his performance is not after all a reason for Ben not to try to construct a powerful moment, even though the spontaneity that would be achieved in so doing really would enhance the musical results. Whatever reasons Ben may have not to try to construct a powerful moment, this is not one of them.[20]

When it is self-defeating for *S* to try to φ for *p*, it need not be necessarily so. It is often the case that *p* is a reason for *S* to φ, but due to contingent features of *S* or the practical context, it will actually impede *S*'s satisfaction of *p* if they set out to φ by being guided by *p*. Consider the distorting effects a job market known to be hyper-competitive can have on the psychologies of entrants to it. A prospective interviewee of course has a reason to give a solid interview, viz., that it could get them job security. But for them to actually try to do so with job security on their

20. Besides the fact that (LR) is textually better supported by what Whiting actually says, the preceding paragraph explains why I reject the recent interpretation of (MC) in Kenneth M Ehrenberg, "How exclusionary reasons guide" (2024) 15:1 Jurisprudence 71. There, he says that (MC) "only requires that reason be *capable* of serving as a motivation" (*ibid* at 71 [emphasis in original]). The problem, as I see it, is that we are capable of being motivated to do what is necessarily self-defeating. For instance, Ben can want to play sax more spontaneously, although, being so moved in his playing, he necessarily frustrates the desire that moves him. On Ehrenberg's view, Ben has a reason to play sax more spontaneously, because the value of spontaneous playing is capable of serving as his motivation, even as being so moved in his playing is necessarily self-defeating. To my mind, this is the wrong result. The intuitive connection between reasons and guidance does not seem to be respected if some reasons are necessarily such that they cannot guide us to success. Thanks to a reviewer for alerting me to Ehrenberg's work.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

mind (how wonderful it would be to be free of applications, to save for the future, to make rent) is liable, the market known to be such as it is, to whip up a state of frenzied over-preparation, which will in the end be detrimental to their giving a solid interview. Although it happens to be self-defeating for the interviewee to try to give a solid interview guided by their job security reason, here it is only *contingently* so. It is not necessarily self-defeating, because there are ways the world could be such that trying is actually conducive to succeeding—as would be the case, for instance, were the job market not known to be in such a dire state.

With the notion of logical possibility clarified, I suspect that many will find principle (LR) uncontroversial. For it is an extremely weak principle, albeit one whose application is often overlooked.[21] It does not rule out that we have reasons for action, even when we are completely ignorant of them, or unmotivated, or indeed when it would be contingently self-defeating to try to act on them. It is also not inconsistent with alternative specifications of (R) that put stronger necessary conditions on when a fact counts as a reason. Given that (LR) is not implausible and in any case very weak, I grant it as an assumption in what follows.[22]

2.2. The Credit Argument

To the extent that Whiting's challenge instantiates the dominant strategy, we can expect there to be an argument as to why the idea that we could act for a reason for a reason is incoherent. Principle (LR) dictates what exactly it is that needs to be argued for here. The argument needs to establish that we cannot act for a reason for a reason in the sense that it is not logically possible for us to do so.

To this end, Whiting offers the following line of thought:

> A plausible view is that, if you φ for a good reason, that is, for an undefeated reason which favours φing, you are creditworthy. . . . Suppose that Kelly decides to go to the pub for the reason that she promised Dave to meet him there. . . . Here, Kelly does what she ought to do and deserves credit for this. Now suppose that Kelly decides to go to the pub for the reason that she promised Dave to meet him there for the reason that she will get a reward if she acts in a creditworthy fashion. That Kelly will be rewarded is a second-order reason—a reason for her to act for certain first-order reasons. However, in this case Kelly is surely not creditworthy for going

21. Moore rejects (LR) as it applies to moral reasons on account of this very weakness: see Moore, *supra* note 6 at 875. He reasons that any specification of principle (R), insofar as, and because, it applies to moral reasons, must be a stronger 'substantive' principle (e.g., grounded in the morality of fairness) rather than a weakly 'formal' one. I do not see why this should be so. Whatever else moral reasons are, they are a species of reasons, and so if the existence of reasons is in general constrained by a formal principle, it seems *ad hoc* to suggest without argument that things should be any different for moral reasons.

22. I do not suggest that (LR) is totally unassailable. For instance, Gardner and Jung have argued that there is a fundamental asymmetry between reasons for and against action, namely, that whilst we can act for reasons for action, we cannot act for reasons against action: see Gardner & Jung, *supra* note 12 at 571. If this asymmetry holds, then assuming that reasons against action are equivalent to reasons for omissions, (LR) could be questioned in its generality across both reasons for and against action.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

> to the pub. . . . The case in which Kelly acts for a reason for the reason that she will get a reward shows that it cannot be true *both* that acting for a good reason suffices for creditworthiness *and* that it is possible to act for a reason for a reason. So, we must reject either the claim about credit or the supposition that in this case Kelly acts for a first-order reason for a second-order reason. . . . I suggest that we reject the latter.[23]

We have here the argument Whiting takes to complete his conceptual challenge. I believe it is faithfully reproduced in the form of a valid *reductio ad absurdum*. We can call it the Credit Argument.

The Credit Argument

P1 When *S* φs for the right reason, they are creditworthy.[24]

P2 It is logically possible for *S* to φ for a reason for a reason.

P3 But when *S* φs for the right reason for a yet further reason, *S* is not creditworthy.

C Thus, when *S* acts for such further reasons, *S* both is, and is not, creditworthy.

Since the conclusion is a contradiction and logically follows from *P1-3*, we are compelled to reject one or more of these premises. Whiting's suggestion is that, faced with these options, we are compelled in particular to reject *P2*. Needless to say, we are compelled to reject *P2* only on the assumption that it is unacceptable to reject *P1* or *P3*. Since the Credit Argument is a valid argument, we will, in the rest of the article, inquire into whether it is sound: whether it can safely be assumed that neither *P1* nor *P3* can be rejected. Sections 3 and 4 focus on *P1*, and section 5 on *P3*.

3. The Right Reasons Are Second-Order

Let us take *P1* first. Notice that the 'right reason', as it figures in that premise, admits of first- and second-order readings. From the perspective of Whiting's challenge, the correct reading must be first-order, else the Credit Argument will assume what it aims to prove incoherent. The problem, which I aim to bring out in this section, is that *P1* is only plausible on its second-order reading. I begin by unpacking the basic thought that *P1* encodes, in virtue of which it is plausible, and then show how it already invokes second-order reasons.

23. Whiting, *supra* note 6 at 406 [emphasis in original, footnotes omitted].

24. I say 'right' reason where Whiting says 'good', although nothing hinges on the choice of words. What matters is that they are both intended, minimally, to describe the reasons for acting for which we are intuitively creditworthy. For the record, I prefer my word for this because it is standard in the literature, including the literature Whiting uses to bolster *P1*. I therefore continue to use it. The narrower reading of 'good' available in the quoted passage, where the 'good' reasons are the undefeated ones, is discussed in detail in section 4.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

There is a powerful strand in moral philosophy whose central thought is that moral worth accrues to those who do the right thing for the right reasons.[25] The plausible intuition driving the thought is that it is in some sense not enough for an action to manifest moral worth that it happens to be the right thing to do. It must moreover be motivated in the right way, responsively to what is morally relevant in the situation at hand. This is the basic thought behind *P1*, and it has an impressive pedigree. In the *Nicomachean Ethics*, for example, Aristotle says that a morally serious person would sacrifice their life for the sake of their friends or country out of clear-eyed pursuit of the noble.[26] Or, perhaps more amenable to the modern ear, there is Kant's suggestion in the *Groundwork* that the shopkeeper who charges a fair price from the motive of duty therein manifests a good will, unlike one who does so merely to preserve their reputation.[27] Let us call the basic thought behind *P1* the 'Aristotle-Kant' thought.

For our purposes, it is helpful to consider the Aristotle-Kant thought as it bears on contexts of choice. We can separate *what* is chosen from *how* what is chosen gets chosen. The emphasis on motivation in the Aristotle-Kant thought concerns less the 'what' than the 'how'. But the proper significance of how we choose what we choose will escape us unless we first understand what it is that we choose from: our options. Suppose a shopkeeper charges £2 for a pack of toilet roll, and that they do so because that is a fair price. Should we say that charging £2 is the option the shopkeeper chose? In one familiar sense, the answer is yes. But that sense, although familiar, is misleading. For in that sense, the shopkeeper chose from an infinity of price options: *n* options to charge £*n*. Of course, for many *n* (where *n* equals 10 billion, say), to charge that price would be quite absurd.[28] To make any money at all, as shopkeepers are wont to do, it would be pointless to charge £10 billion for a pack of toilet roll. In a less misleading sense of 'option', that is not really an option. This makes it plausible that our options are not just actions, but actions done for reasons: to charge £2 because it is fair, or £5 because it capitalises on a supply shortage, and so on.[29] In

25. This is not to say that its representatives agree on the smaller details of what the thought amounts to. See especially Nomy Arpaly, "Moral Worth" (2002) 99:5 J Philosophy 223; Julia Markovits, "Acting for the Right Reasons" (2010) 119:2 Philosophical Rev 201. It should be noted that Whiting acknowledges the strength of this literature, and indeed seems to take this all by itself to push the rejection of *P1* beyond the pale: see Whiting, *supra* note 6 at 406. I agree that this is an excellent literature, and that it lends serious support to *P1*. As we will see, however, I do not agree with Whiting that it necessarily supports *P1* on his preferred first-order reading.

26. See Aristotle, "Nicomachean Ethics" in Jonathan Barnes & Anthony Kenny, eds, *Aristotle's Ethics: Writings from the Complete Works*, translated by WD Ross (Princeton University Press, 2014) 207 at 339 (1169a 18-22).

27. See Immanuel Kant, "Groundwork of the metaphysics of morals" in *Practical Philosophy*, edited & translated by Mary J Gregor (Cambridge University Press, 1996) 37 at 52-54 (4:385 at 4:397-98).

28. Hence, Raz lambastes as "silly" the suggestion "that as I write I have a choice not only between carrying on with this chapter or taking a break to listen to the news, but also between either of these options and cutting off my finger or my ear, or taking my shoes out of the cupboard and putting them back in again, etc." Raz, *supra* note 13 at 178.

29. What about those cases some believe to be felicitously described as cases of choosing an option for literally no reason, such as Chang's nice example of choosing one can of soup from an aisle

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Christine Korsgaard's colourful phrase, we choose "the whole package."[30] What we choose when we choose to do the 'right thing' is, like any option, a whole package: an action done for a certain reason. The significance of how we choose what we choose lies in what it reveals about the value we see in the relevant whole package, the value we see in doing a certain thing for a certain reason.[31]

In the previous paragraph, I claimed that reasons individuate the objects of choice. This is a common, although not uncontroversial, view. Let me therefore pause for a moment to defend this claim, before returning to my mainline argument.[32] One source of the controversy is metaethical. The worry is that what drives the claim that reasons individuate the objects of choice, so that our options are actions-for-reasons, is the desire to individuate the morally right option in terms of the morally worthy kind of reason for which it is taken. In other words, the worry is that the claim is bound up with a Kant-inspired, if not Kantian, conflation of deontic normativity—what the categorical imperative says to do, which it is impermissible or wrong not to do—and the aretaic normativity of moral worth.[33] But even if the Kant-inspired view entails the claim in question, the claim does not entail the Kant-inspired view. There is no inconsistency in believing, as I do, that our reasons individuate the right thing to do, whilst also believing that we do not have to do the right thing for the right reasons, however morally worthy it may be to do so.

---

of identical cans of soup? See Ruth Chang, "Can Desires Provide Reasons for Action?" in R Jay Wallace et al, eds, *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (Oxford University Press, 2004) 56 at 82. I must admit to not finding it entirely clear how these cases are meant to work. At one level of description, the can *is* chosen for the same reason that any can would have been, viz., that one is responsible for dinner tonight, or something like that. If it is said instead that the object of choice is not *this* can, but *this-rather-than-that* can, then it is not clear to me that the context should be described as one of choice. Or no more so than, say, my 'choice' this morning to shave in whatever fine-grained pattern of strokes I did, as opposed to any other (suitably non-eccentric) pattern.

30. Christine M Korsgaard, "Acting for a Reason" in *The Constitution of Agency: Essays on Practical Reason and Moral Psychology* (Oxford University Press, 2008) 207 at 218. Korsgaard is not unique in this view. In a prescient paper, Mary Glover endorses it thus: "Is it right to plunge a knife into the body of a living man? The answer would appear to be 'No'; yet it is sometimes right to remove his appendix, and this cannot be done without first cutting him open. What distinguishes surgery from murder . . . is precisely the intention." Mary Glover, "Obligation and Value" (1938) 49:1 Ethics 68 at 69. What the surgeon and the murderer choose to do differs, because objects of choice are individuated not just by actions but the ends pursued in doing them. Barbara Herman puts the point in a Kantian idiom when she says: "[T]he content of a maxim of action must be the described action and its end. . . . Actions and ends are chosen as they are seen to be good . . . . The good-making characteristics of the action-end pair, its principle, must therefore be a constitutive element of the maxim's moral content." Barbara Herman, *The Practice of Moral Judgment* (Harvard University Press, 1993) at 223. That is, objects of choice are individuated not just by actions but by ends. Thanks to a reviewer for pointing me to friends of Korsgaard on this point.

31. Or, in Herman's words: "[W]e must look to the agent's motives if we are to describe the action-as-willed. If the end of an action is the state of affairs the agent would bring about, the motive explains what it is in the state of affairs that is attractive. We do not see what the agent wills until we see her action and end in the way she sees them as good." Herman, *supra* note 30 at 221 [footnote omitted].

32. I am very grateful to a reviewer for urging me to discuss this claim in more detail.

33. See e.g. TM Scanlon, *Moral Dimensions: Permissibility, Meaning, Blame* (Harvard University Press, 2008) at 59-60; Audi, *supra* note 2 at 62.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Another source of the controversy is action-theoretic. The worry is that if our options are actions-for-reasons, and we can pursue options for reasons, then we will have to countenance the higher-order option of pursuing options-for-reasons, and with it the possibility of yet higher-order reasons for pursuing options-for-reasons. In that case, for any n^{th} order option we countenance, there will always be an $n+1^{th}$ order reason for pursuing it. This becomes an issue if we assume that (a) our options are intentional actions, (b) our actions are only intentional if we comprehend the reasons for which we do them, and (c) we have finite comprehensional capacities. On these assumptions, for any n^{th} order option we countenance, there will always be an $n+1^{th}$ order reason for pursuing it that we must be able to comprehend—an ability which, as n increases, it becomes increasingly certain we do not possess. To avoid this issue, we must either roll back on the controversial claim, or explain which of (a), (b), and (c) should be denied. Since (a) and (c) are highly plausible, (b) is the natural candidate. But some, such as Robert Audi, believe it is hopeless to deny (b), because doing so requires a view on which "the idea of intentional action as action for a reason would not apply."[34] On closer inspection, however, this does not seem to be required. All that is required is a view on which acting for a reason does not entail comprehending the reason for which the act is done. And we can, I think, be guided by reasons without comprehending what it is about the situation that gives us the reason that guides us. The classic example is Huck Finn, whose explicit deliberations indicate that he does not comprehend why it would be wrong to hand Jim over to the slave hunters, and yet, when the time comes, conceals Jim from the slave hunters and thereby does the right thing. Huck Finn's lying to the slave hunters is intentional, because it is guided by a reason, despite incomprehension of what really constitutes the reason that guides him.[35]

A final source of controversy is psychological. The worry is that if our options are actions-for-reasons, and we can pursue options for reasons, then that would imply that we can act-for-a-reason for a reason, which is not in fact psychologically possible for us.[36] I agree with Scanlon and Audi that, as a psychological matter, we are unable to simply choose the reasons for which we act, and *a fortiori* the reasons for which we act-for-a-reason. But even if, in this sense, we lack 'direct' control over the reasons for which we act, it is not obvious in this context why it should matter over and above more 'indirect' forms of control. I therefore struggle to see why only 'direct' control would be enough to furnish the psychological possibility of acting-for-a-reason for a reason. In this connection, Mary

34. Audi, *supra* note 2 at 64, n 15.

35. See Mark Twain, *The Adventures of Huckleberry Finn* (Sterling, 2006). Later on, I suggest that moral worth is a matter of acting for a certain type of second-order reason. In allowing that we can be guided by reasons *de re*, acts that come out morally worthy on my view need not be prone to the charge of fetishising morality, i.e., doing the right thing, not responsively to what matters, but responsively to the thought of doing the right thing—a thought Bernard Williams famously describes as "one thought too many." Bernard Williams, "Persons, Character and Morality" in *Moral Luck: Philosophical Papers 1973-1980* (Cambridge University Press, 1981) 1 at 18

36. See Scanlon, *supra* note 33 at 59-61; Audi, *supra* note 2 at 65.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Glover, for one, observes of our motivations and emotions that "we all habitually control them to some extent; indirectly, through the direction of our attention," and her examples of such "indirect control" are both familiar and psychologically plausible.[37]

> We can inflame anger by brooding upon it; on the other hand, we can avoid being made angry by the sufferings of other people by refusing to attend to them; [and] we can avoid the joy of elevated thoughts by never thinking about elevated things.[38]

Even Audi, who regards only 'direct' control as sufficient, concedes that Glover's observations are both "important and empirically plausible."[39]

Of course, whether any form of control is psychologically possible for us is, as I mentioned before, an empirical question to be settled by science. Giuseppe Rocché, in an illuminating discussion, bases his argument against exclusionary reasons on an empirical study that he takes to suggest the psychological impossibility of refraining from acting-for-a-reason for a reason, that is, of excluding reasons for doing otherwise in this way.[40] In the study, researchers sought to test the efficacy, in judges' deliberative processes, of a statutory rule that prohibits the use of evidence concerning the sexual history of alleged victims in sexual assault trials.[41] Judges were presented with a sexual assault case and asked whether they would convict the defendant. The control group was given no further evidence, whilst the 'suppression' group was given further evidence, the admissibility of which, under the statutory rule, was up to them to decide. If judges are given the further evidence and find it inadmissible under the rule, then if they are psychologically capable of excluding it, we would expect 49.1% to convict, as in the control group which simply lacks the further evidence. But if the same judges cannot exclude it, we would expect only 7.7% to convict, like those in the suppression group who receive the further evidence but find it admissible. The study found that judges who receive the further evidence and find it inadmissible

37. Glover, *supra* note 30 at 73.
38. *Ibid.*
39. Robert Audi, "On Mary Glover's 'Obligation and Value'" (2015) 125:2 Ethics 525 at 528. These examples of 'indirect' control may seem more amenable to the wrong side of what Ehrenberg calls Raz's "dispute with HLA Hart over the distinction between exclusionary reasons as Raz understood them and peremptory reasons as Hart understood them." Ehrenberg, *supra* note 20 at 73. But that is not so. For if we have 'indirect' control in that we are psychologically capable of simply leaving certain reasons out of consideration in practical deliberation, we thereby have the same sort of control over whether we act for those reasons. Thus, even if, as Ehrenberg says that normatively speaking, "exclusionary reasons exclude acting on some reasons … while peremptory reasons prevent deliberation on certain reasons" psychologically speaking, an ability to prevent deliberation on certain reasons is itself a way of excluding acting for those reasons (*ibid* at 74).
40. See Giuseppe Rocché, "Exclusionary Reasons and Mental Contamination" (2024) 52:1 Revus 1 at 15. Many thanks to a reviewer for alerting me to this interesting cross-disciplinary literature, and for pressing me to discuss the psychological possibility of acting for second-order reasons in finer detail, including the bearing of empirical studies on the view I am developing.
41. See Andrew J Wistrich, Chris Guthrie & Jeffrey J Rachlinski, "Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding" (2005) 153:4 U Pa L Rev 1251 at 1298-1303.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

convict at a rate of 20%, indicating that many of them were in fact psychologically unable to exclude evidence despite regarding it as inadmissible under statute.[42]

This is important and interesting, but does not establish what it needs to. All it establishes is that, whilst most judges (and so perhaps most of us) are psychologically unable to exclude considerations in deliberation, some *are* able to do so. Thus, more than anything, the study suggests that refraining from acting-for-a-reason (i.e., deciding based on inadmissible evidence) for a reason (i.e., the statutory rule) is psychologically hard to do, and that not everyone—fancy judges included—has the mental fortitude to do so. Chances are, there exist other studies that bear on the psychological possibility of acting-for-a-reason for a reason of which I am not aware, and these should be considered on a case-by-case basis. But whilst I cannot, of course, deny that knock-down empirical evidence would threaten the present argument, I do wonder what sort of study could provide such evidence, especially given the intuitive plausibility of forms of 'indirect' control. Even if, for example, it so happens that 0% of subjects in a given study are psychologically able to exclude reasons, on whatever operationalisation, it is not clear why that would prove it psychologically impossible to do, rather than simply too hard for all tested subjects.[43] For now, then, let us grant the psychological possibility of acting-for-a-reason for a reason, and with it the claim that reasons individuate our options.

This claim, recall, was needed to explain the emphasis on motivation in the Aristotle-Kant thought, which is the basic thought behind *P1* of the Credit Argument. The thought, roughly, is that the significance of the 'how' of choice lies in what it reveals about the value we see in 'what' option is chosen, considered as what Korsgaard would call a "whole package"—that is, the value we see in doing a certain thing for a certain reason.[44] Put this way, it is clear that the Aristotle-Kant thought is consistent with different substantive views on the content of the value we see. It is in this sense a metaethical thought. The content of this thought, I submit, is that moral worth is a matter of acting for a certain type of (positive) second-order reason. To see which type, we need to consider the ways in which, when there is value in doing something—call this the 'basic' value—there may in turn be *further* value in doing it guided by recognition of its basic value. Since there being basic value in doing something is a reason, p, to do it, the ways there may be further value in doing it guided by its basic value are so many

42. See *ibid* at 1302.

43. It is worth mentioning, as Ehrenberg does well to show, that excluding certain reasons may not be the only way of being guided by exclusionary reasons: see Ehrenberg, *supra* note 20 at 72. For it may be that, when we act against the valence of an exclusionary reason for a non-excluded reason, we only allow ourselves to act for the non-excluded reason (as opposed to any excluded reason) because the exclusionary reason so guides us. In that case, even proof that we are psychologically unable to exclude reasons would still leave it open whether we are, more generally, psychologically capable of guidance by exclusionary reasons.

44. Korsgaard, *supra* note 30 at 218.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Figure 1. Proposed typology of second-order reasons.

In what relation do the basic and further values stand?		*In what relation does acting from recognition of the basic value stand to (precluding) the realisation of the further value?*	
		Contingent	Constitutive
	Sameness	TYPE A	TYPE C
	Difference	TYPE B	TYPE D

ways there can be reasons, q, to do it for the reason p, i.e., so many types of second-order reason.[45]

The Figure 1 table may be helpful as a visual aid to the proposed typology. There are four axiological scenarios to consider, which vary along two dimensions: the relations between the basic value and the further value, and the relations between acting from recognition of the basic value and (precluding) the realisation of the further value. There are, correspondingly, four types of second-order reason.[46]

In Type A cases, the further value is the very same value as the basic value. For example, if Fred needs food, then there is value in his going to the shop insofar as this satisfies his need, which gives him, other things equal, a reason to go to the shop. Maybe he ends up there because he goes past it whilst jogging, or because he has to collect a package, or because he ducks in to avoid the rain. In each case, Fred goes to the shop, thus satisfying the reason he has to go to the shop in virtue of his need. Equally, he could go because he sees that doing so has the value of satisfying his need.[47] Like any of the ways that Fred could end up at the shop, ending up there guided by his need for food would of course satisfy that very need. That is, the further value of going to the shop guided by the basic value in doing so is none other than the basic value. As we saw in

45. For the view that the paradigm case of a reason for action is the value of the action it is for, see Ulrike Heuer, "Value-Based Reasons" in Russ Schafer-Landau, ed, *Oxford Studies in Metaethics Volume 20* (Oxford University Press, 2025) 224. I find Heuer's view compelling, but it is, of course, contested ground. As far as I can tell, however, the typology that follows is valid independently of our views on the relation between reasons and values. Any formulations that suggest otherwise should be taken as for ease of exposition only.

46. As mentioned in the text accompanying *supra* note 12, I include omissions amongst actions. Thus, if there is basic value in doing something, then the further value there may be in doing it guided by recognition of its basic value can equally be the further value in omitting or refraining from doing it guided by recognition of its basic value. That is, the typology includes negative, as well as positive, second-order reasons. In the negative case, the basic and further values remain either the same or different; but crucially, unlike the positive case, the further value is such that acting from recognition of the basic value *precludes* its realisation either necessarily or contingently. Note that each type in the proposed typology admits of positive and negative valence, which, if we allow that to define types, results in eight types of second-order reason. For ease of exposition, I will focus on the positive case, since that is what the Credit Argument makes salient.

47. In which case, the value of need-satisfaction not only gives Fred a reason to go to the shop but a reason to do so for that reason, or as Raz might say, both a reason for conformity and a reason for compliance: see Raz, *supra* note 3 at 179. To avoid needless introduction of terms, I refrain from using the Razian language in what follows.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

subsection 2.1, it can be a contingent matter whether or not the basic value of a given course of action also doubles as a further value. In Fred's case, it does—just so long as his going to the shop for his need-based reason remains an effective way of going to the shop. But if this ceases to be, then it becomes contingently self-defeating for Fred to go to the shop for his need-based reason.[48] In that case, if there remains any further value in Fred's going to the shop guided by the basic value in doing so (viz., need-satisfaction), it cannot be the same as this basic value.

Type B cases differ in that the further value is a different value from the basic value. For example, if it would please Fred's therapist were he to make an effort with his day-to-day, and his going to the shop guided by the value in doing so (viz., need-satisfaction) would prove this, then on the one hand, that it satisfies his need is a reason for him to go to the shop, and on the other, that his doing so for that reason would please his therapist is a reason for him to do so for that reason. This is a second-order reason to go to the shop for the need-based reason, but it differs from—and competes with—the Type A reason Fred also possesses. We know it differs because the further value that underwrites it is not the value of satisfying Fred's need but the quite different value of pleasing his therapist. The basic and further values come apart. As with Type A cases, the further value in Type B cases provides a reason to act guided by the basic value only insofar as doing so is an effective way to realise the further value, which is a status it has contingently. Perhaps what pleases Fred's therapist depends on her mood: Sometimes she advises that being kind to oneself consists in upholding day-to-day tasks, and other times that it requires relaxing the demands one makes of oneself day-to-day. In the latter moods, Fred's going to the shop for the need-based reason will no longer please his therapist, although it will, other things equal, still facilitate the satisfaction of his need for food. Fred's going to the shop for the need-based reason will then only be supported by the Type A second-order reason, not the Type B one.

What unites cases of Types A and B is that the further values can be realised independently of whether the protagonists act from a recognition of the basic values. Hence why, in these cases, acting from such recognition may or may not realise the further value, contingent upon whether or not so doing actually stands in the appropriate instrumental relation to realising the further value. When it does, the second-order reason there is to do so is a *contingent* one. This contrasts with Type C and D cases, where the further value cannot exist independently of action guided by the basic value. In these cases, the relation between acting guided by the basic value and realising the further value is *constitutive* rather than contingent. The difference between Type C and D cases turns on whether or not

48. Situations where the basic value in someone's doing something does not double as a further value in their doing it guided by the basic value—such that doing the latter is contingently self-defeating—are a precondition for a Razian authority justifiably to step in according to his 'normal justification thesis'. See Joseph Raz, *The Morality of Freedom* (Oxford University Press, 1986) at 53-57.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

the basic and further values are one and the same. Plausibly, if there can be *sui generis* value in declarations of true love, then realising it requires satisfying a Type C constitutive second-order reason. The *sui generis* value, we are assuming, cannot be realised without the declaration, but there is no such declaration unless it is made for the reason that one truly loves its recipient. Similar forms of words or gestures produced for any other reason cannot but fail to realise the *sui generis* value, which is what it means to say they are empty. With Type D cases, by contrast, the basic and further values are different (as in Type B cases), but still the further value cannot be realised except by acting from recognition of the basic value in doing so (as in Type C cases). We will consider examples of Type D cases in a moment. For now, suffice it to say that the further value in both Type C and D cases cannot be realised without guidance by the basic value, which explains why they provide constitutive positive second-order reasons.

Typical examples of morally worthy behaviour marshalled by proponents of the Aristotle-Kant thought emphasise their Type D character, often by contrasting them with Type B variants where moral worth is intuitively lacking. Consider Julia Markovits' cases of the reward-seeker and the altruist.[49] Both rush into a burning building from a recognition of the fact that doing so enables the rescue of a trapped child. But they each see a different further value in doing so, guided by that same basic value. For the reward-seeker, the further value is that she will thereby earn a reward. This value accrues to her rushing into the building guided by the basic value of the child's life only insofar as a reward happens to be set up for her doing so, as would be the case, for instance, if a millionaire with insight into the reasons for which she acts offers her such a reward. Here the further value underwrites only a Type B contingent second-order reason. For the altruist, by contrast, the further value lies in due concern for the intrinsic value of the child's life, or something like that. Rushing into the building in order to rescue the child is constitutive of what will manifest such concern, such that this value, given the situation and its concrete demands, simply cannot be realised in any other way. The further value the altruist sees supplies a Type D constitutive second-order reason. Needless to say, it is the altruist and not the reward-seeker whose action is supposed to be morally worthy.

A similar structure can be discerned in Nomy Arpaly's cases of Ron the extremist and Huck Finn.[50] Ron wants to kill Tamara, but refrains so as to realise the value (read: 'value') of not killing members of his own ethnic group, to which Tamara also belongs. It is clear that Ron's acting, guided by this basic 'value', stands only contingently related to the further value of protecting Tamara's life. Had the world been set up differently—if Tamara happened, say, not to share Ron's ethnicity—then Ron's acting from the basic 'value' would no longer be an effective way of preserving Tamara's life. We thus have a contingent second-order reason of Type B. By contrast, when Huck lies to the slave hunters in order to save Jim, he does so because he sees that this is what the further value

49. See Markovits, *supra* note 25 at 227-28.
50. See Arpaly, *supra* note 25 at 227-30.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

of due respect for Jim's humanity demands, however viscerally, inexplicitly, and against his own conscience. Upon being confronted by the slave hunters, the situation's concrete demands leave Huck without a way to realise this value other than by lying to them in order to save Jim. Here the further value supplies a constitutive second-order reason of Type D. Although Ron and Huck both do the right thing, it is Huck alone that Arpaly wants to suggest acts in a morally worthy way.

For our purposes, Markovits, Arpaly, and other proponents of the Aristotle-Kant thought need not share any substantive view on the further values that determine moral worth in their preferred examples, nor need they be happy with our way of parsing the examples out. All that matters is that their common Type D structure can be discerned. What unites proponents of the Aristotle-Kant thought is that the further value in their preferred examples, whatever its content, supplies *constitutive positive second-order reasons* of *Type D*. Since, as I have suggested, *P1* is plausible only insofar as it expresses this distinctive metaethical thought, the 'right reasons', I claim, should be given a second-order reading.

4. The Right Reasons Are Not First-Order

In the previous section, I suggested that the sorts of examples typically used to support *P1* are best described in second-order terms, as cases of Type D. But that suggestion is, after all, just a suggestion, and I have not tried to argue that the examples cannot be described in first-order terms. So, perhaps charity demands that we give the 'right reasons' in *P1* a first-order reading. The most sensible proposal along these lines is probably Whiting's own view that the first-order reasons—acting for which is creditworthy—are the *undefeated* reasons.[51] Is there more to be said for the Credit Argument on this reading?

I do not think so. Acting for an undefeated reason is simply not sufficient to make the action creditworthy. This can be established with counterexamples involving very mundane contexts of action, where whatever we do, there is nothing especially significant at stake. Take the choice at the breakfast bar between apple juice and orange juice. This is probably a choice between incommensurables, such that neither reason for drinking a juice is defeated by the other: Both are undefeated. But it is implausible that we would therefore deserve credit for drinking either juice for the reason in its favour. This is partly due to the implication that we deserve credit whatever we do, which we may, and not implausibly, think is analytically ruled out by the very notion of creditworthiness. More basically, however, what is implausible is that we could deserve credit for something so utterly mundane, just because we do it for an undefeated reason. We can see this by tweaking the incommensurability out of the case, so that there is a clearly superior option, for instance, by making the choice one between a normal glass of apple juice and a tiny one. Other things equal, our reason to opt for the

51. See Whiting, s*upra* note 6 at 406.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

tiny glass is defeated by our reason to opt for the normal one, so that only the latter reason is undefeated. But even if we opt for the normal glass guided by the undefeated reason, so that we are justified in doing so, it is again implausible that we are thereby creditworthy for so doing. As John Gardner says, "more is required for credit than is required for mere justification."[52]

This shows that for *P1* to come out true with the 'right reasons' read as undefeated reasons, it needs to be modified. The difficulty lies in modifying it such that it comes out true, whilst retaining the link it seeks to establish between the right reasons and creditworthiness. The only real candidate seems to be something like this:

> *P1** When *S* φs for the right reason (i.e., an undefeated reason), they may be, but need not be, creditworthy.

This is probably the strongest true proposition that connects acting for the right reason and creditworthiness, reading the 'right reasons' as undefeated reasons. We could arrive at *P1** by observing that, whilst it is not plausible that acting for an undefeated reason is sufficient for that action's being creditworthy, it is plausible that an action's being creditworthy is sufficient for its being done for an undefeated reason. So when someone acts for an undefeated reason, it is perfectly possible that they deserve credit for this. That is what *P1** captures.

Modifying *P1* to *P1** may well restore soundness to the Credit Argument, but it does so by blocking its desired conclusion. If, as *P1** says, *S* may, but need not, be creditworthy in φing for the right reason, then there may, but need not, follow a contradiction in the presence of the other premises. For only some of the second-order reasons mentioned in *P3* will generate a contradiction: those that bear on acting for whichever 'right reasons' it is creditworthy to act. Only *those* second-order reasons turn out to be incoherent. But that is a much narrower claim than Whiting hoped to establish. The desired conclusion is that second-order reasons are *all of them* incoherent, but on the modified Credit Argument, some of them turn out not to be.

5. Do the Examples Generalise?

Some readers may already agree with my section 3 contention that the 'right reasons' in *P1* are best interpreted in second-order terms. Others may disagree with me on that point, but agree with my section 4 claim that reading the 'right reasons' as undefeated reasons is no good. Some of these may already regard this as pushing the first-order reading of *P1* beyond the pale, and with it, the only

52. John Gardner, "Justifications and Reasons" in *Offences and Defences: Selected Essays in the Philosophy of Criminal Law* (Oxford University Press, 2007) 91 at 104. I remain silent on the question whether it is also the case that *less* is required, i.e., whether it is not only not sufficient but also not necessary for creditworthiness that an action be justified. If that is so, it favours the view that, unlike justification, people are principally creditworthy for something other than what they do (e.g., what they intend to do, who they are, etc.).

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

reading of *P1* they were prepared to accept. But others may be willing to hold out for *some* more promising first-order reading. After all, nothing in the intuitive resonance of the Aristotle-Kant thought makes it such that the failure of the undefeated reasons reading forces us to accept a second-order one. So, let us grant that some first-order reading of *P1* is acceptable. In this final section, I argue that this does not help the Credit Argument, since *P3* is false.

P3 says: When *S* φs for the right reason for a yet further reason, *S* is not creditworthy. One way to bring out the falsity of *P3* is to point out that it is based on examples from Whiting the features of which simply do not generalise. The first of these we have already seen:

> [S]uppose that Kelly decides to go to the pub for the reason that she promised Dave to meet him there for the reason that she will get a reward if she acts in a creditworthy fashion. That Kelly will be rewarded is a second-order reason—a reason for her to act for certain first-order reasons. However, in this case Kelly is surely not creditworthy for going to the pub.[53]

Call this example *Pub*. In *Pub*, the fact that Kelly promised Dave to meet him at the pub is a first-order reason for her to go to the pub, and the fact that Kelly will get a reward if she acts in a creditworthy way is a second-order reason for her to go to the pub for the first-order reason that she promised Dave to meet him there. Incidentally, Whiting seems aware of the risk that his examples might not generalise, since he offers the second example to convince the reader "that nothing turns on the particular example."[54]

Whiting's second example is as follows

> Kelly is deciding whether to send her daughter to school A or school B. The career-related considerations favour B, but the weightier education-related considerations favour A. Kelly promised Dave that she would make her decision on educational grounds alone. So, Kelly decides to send her daughter to A for educational reasons on the basis of her promise to Dave. As it happens, had Kelly not made her promise, she would have sent her daughter to B, not A. She is concerned about her daughter's education only because she is concerned with keeping her promise to Dave. That is, she responds to the education-related considerations only for the reason that she promised Dave to do so.[55]

---

53. Whiting, *supra* note 6 at 406.

54. *Ibid.* It may also be that, if Kelly acts to get a reward for acting in a creditworthy fashion, she necessarily fails to act in a creditworthy fashion, and so necessarily cannot be rewarded for so acting. That is, it may, in my terms, be necessarily self-defeating for Kelly to try to satisfy her second-order reason *by following it*, so that it is not logically possible for her to satisfy it in this way. In that case, *Pub* cannot show anything about the creditworthiness or otherwise of acting for the right reason for a further reason, because the putative further reason fails principle (LR) and is thus in fact no reason at all. This may indeed be what is going on in *Pub*, although note that the argument of this section applies even if it is not. Thanks to a reviewer for alerting me to this additional weakness in *Pub*.

55. *Ibid* at 406-07. Whiting imagines the education-related reasons to be "That the school is nearby" and "that it offers philosophy," and the career-related reasons to be "that having a child there would further Kelly's career" (*ibid* at 399).

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Call this example *School*. There are various distracting factors at play in *School*, but its core content is as follows. That her daughter will receive a better education at the educationally better school is a first-order reason, an education-related reason, for Kelly to send her daughter there (in this case, school A). That she promised Dave to make her decision only on the basis of education-related reasons is a second-order reason for Kelly to send her daughter to school A for a certain first-order reason, namely, the education-related reason.

The first-order reasons in *Pub* and *School* are meant to strike us as the 'right reasons', if any are. Let us grant that these first-order reasons are indeed the right reasons. In that case, *Pub* and *School* really do bear out the truth of *P3*. After all, in both examples, Kelly is described as acting for the right reason for a further reason, whilst our intuition tells us that she is not creditworthy for acting as she does. Nevertheless, I submit, we can agree on intuitions around *Pub* and *School* without swallowing any general claim about the creditworthiness or otherwise of acting for the right reason for a further reason, such as is expressed in *P3*. The intuition that Kelly is not creditworthy for acting for the right reason in *Pub* and *School* can be attributed to features of the examples other than that she does so for a second-order reason. Consider the following variants of *Pub* and *School*.

> *Pub** Kelly goes to the pub for the reason that she promised Dave to meet him there, out of respect for the practice of promising.

> *School** Kelly decides to send her daughter to the educationally better school for the reason that she will receive a better education there, out of love for her daughter.

In *Pub**, Kelly has the same first-order reason to go to the pub as in *Pub*, viz., that she promised Dave to meet him there. But her second-order reason to go to the pub for this first-order reason is a different one, and could be rendered as the fact that going to the pub for the reason that she promised Dave to meet him there is what would constitute the value of respect for the practice of promising in the circumstances at hand.[56] Similarly, in *School**, Kelly has the same first-order reason to send her daughter to the educationally better school as in *School*, viz., that she will receive a better education there. But again, her second-order reason to send her daughter to the educationally better school for this first-order reason is a different one. The relevant consideration is that sending her daughter to the educationally better school for education-related reasons is what properly constitutes her love for her daughter.

56. On a Razian account, the fact that Kelly promised Dave is a first-order reason to go to the pub, protected by an exclusionary reason not to act for certain other first-order reasons not to go to the pub (e.g., that it would be inconvenient). Thus, if Kelly goes to the pub because she promised Dave, she goes for the first-order reason that she promised, not the negative second-order reason. Since in *Pub** the value of respect is constituted by Kelly's going to the pub for the first-order reason that she promised, giving her a positive second-order reason, there is no additional complexity generated by the fact that the first-order promissory reason happens also to be protected by a negative second-order reason. Thanks to a reviewer for pushing me to clarify this.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

In terms of the typology of section 3, the basic difference between the two pairs of examples is that the positive second-order reasons in *Pub* and *School* are plainly *contingent*, whereas in *Pub** and *School** they are more plausibly *constitutive*. In *Pub* and *School*, it is a contingent matter that the further values accrue to Kelly's acting for the first-order reasons. Were the world instead set up so that the reward in *Pub* were not earnable by Kelly's acting for her promissory reason, or so that her promise to Dave in *School* were not fulfillable by acting for her education-related reason, then acting for these first-order reasons would stop being an effective way to realise the relevant further values, and the second-order reasons they underwrite would evaporate. But in *Pub**, it seems that, given her promise to meet Dave at the pub, it is constitutive of what it takes for Kelly to properly respect the promising practice that she meet Dave at the pub for the reason that she promised.[57] Similarly, in *School** it seems that, now that her daughter is about to start school, Kelly cannot, other things equal, make good on her love for her daughter except by making the schooling decision for education-related reasons alone.[58] If my proposed typology is anything to go by, then it should be no surprise that Kelly is intuitively creditworthy in *Pub** and *School**, but not in *Pub* and *School*.

Whatever the merits of the typology, the key point for the present argument is that *Pub** and *School** differ from their unstarred counterparts only in terms of the second-order reasons. The first-order reasons are untouched, and presumably, therefore, remain the 'right reasons'. Like their unstarred counterparts, then, *Pub** and *School** are cases where Kelly acts for the right reason for a further reason. Yet unlike in their unstarred counterparts, Kelly is intuitively creditworthy in *Pub** and *School**. So, if there is an intuition that Kelly is not creditworthy in the unstarred cases, what their starred counterparts show is that it stems from some feature other than that they involve acting for the right reason for a yet further reason. Someone's acting for the right reason for a further reason is insufficient to rule out their being creditworthy in so doing. *P3*, in other words, is false.

Faced with cases like *Pub** and *School**, defenders of the Credit Argument could try to amend *P3* to accommodate them in the following, somewhat *ad hoc*, sort of way.

> *P3** When *S* φs for the right reason for an *irrelevant* further reason, *S* is not creditworthy.

The idea with *P3** is to stipulatively define the 'irrelevant' further reasons as the class of second-order reason found in *Pub* and *School*, and similar such cases,

57. There can be other ways to show respect for the promising practice (e.g., by running a campaign or making a documentary), but the point is that, for Kelly, given her promise to Dave, she now cannot do so other than by doing as she promised for the reason that she promised.

58. Other things might not be equal when there are serious enough reasons to make the schooling decision on grounds other than education, such as that it will massively uproot family life, e.g., because attending the educationally better school will require a house move, or require Kelly to make prohibitively lengthy daily school runs, and so on. Thanks to a reviewer for this point.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

whatever class that may be. Perhaps it is extensionally equivalent to the class of contingent positive second-order reasons found under Type A, Type B, or both—although it need not be. Cases like *Pub** and *School** would not be counterexamples to *P3**, since the protagonists in those cases, although creditworthy, are stipulated to act for the right reason for 'relevant' further reasons. So, unlike *P3*, *P3** does not claim that the protagonists in *Pub** and *School** are not creditworthy. Such cases give us no reason to think *P3** is false.

Once again, the problem with this manoeuvre is that the contradiction it generates will be limited in scope as per the stipulation: It is only when people act for *irrelevant* further reasons that they come out both creditworthy and not creditworthy. In turn, the Credit Argument will only establish that *some* second-order reasons are incoherent, namely, the irrelevant ones. And this is, once again, a much narrower claim than desired. Whiting's challenge set out to show second-order reasons incoherent *tout court*, but it now admits that, in fact, some second-order reasons *are* conceptually coherent: the *relevant* ones.

6. Conclusion

The Credit Argument aims to expose second-order reasons as incoherent. I proposed a novel typology of second-order reasons, and suggested that *P1* of the Credit Argument is plausible only if we read the 'right reason' as a certain type of (positive) second-order reason. I then argued that on the alternative first-order reading, *P1* comes out false, a fault that can only be rectified by seriously weakening the argument. Finally, I argued that even if some other first-order reading turns out to be workable, this will not save the Credit Argument, because *P3* is anyway false, and is, again, not fixable without seriously weakening the argument.

I conclude that the Credit Argument does not go through, and that Whiting's conceptual challenge collapses with it. As ever, important questions remain, such as those surrounding the distinction between constitutive positive second-order reasons of Types C and D, how well the typology accommodates intuitions in examples of negative second-order (i.e., exclusionary) reasons, and whether there are other ways of pursuing the dominant strategy that merit further study. For now, I would be happy to have said enough to show that Whiting's version is unsuccessful. In so doing, I hope to have rendered the concept of second-order reasons more serviceable to those of us attracted to explaining normative phenomena in its terms.

Acknowledgements: I would like to thank Filippa Ronquist, Harry Long, and three reviewers at the Canadian Journal of Law & Jurisprudence for helpful comments on various versions of this article. Special thanks are due to Daniel Whiting for his warm and encouraging feedback, and to Ulrike Heuer, whose thoughts and feedback have had their influence at every stage. This research was made possible by the financial support of the AHRC (grant reference number: 2387877).

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press

Declaration: Competing interests: The author(s) declare none.

Kenta Sekine is a researcher and philosopher. His academic research lies broadly across moral, legal, and political theory, and he has published work on the normative foundations of contract in the liberal tradition. He is a policy researcher at Citizens Advice in the Consumer & Public Services team. Views expressed in this article are his own. Email: kentasekine1@gmail.com.

https://doi.org/10.1017/cjlj.2026.10062 Published online by Cambridge University Press