

Love Redirected: On Adam Smith's Love of Praiseworthiness

Published as: Sivertsen, Sveinung Sundfør (2017). Love Redirected: On Adam Smith's Love of Praiseworthiness. *Journal of Scottish Philosophy* 15 (1):101-123. DOI:10.3366/jsp.2017.0154

1 Introduction

Sweet princes, what I did, I did in honour,
Led by the impartial conduct of my soul:
And never shall you see that I will beg
A ragged and forestall'd remission.
(Henry IV Part 2 Act 5 Scene 2)

If Adam Smith had wanted a literary figure to illustrate his *Theory of Moral Sentiments* (TMS, 2002),¹ he could have done much worse than to pick Shakespeare's Lord Chief Justice. In the course of a single scene towards the end of *Henry IV Part 2*, this figure plays out most of the major themes of Smith's moral theory.

In the scene in question, Henry IV has just died, and his son, Prince Hal, is set to succeed him. Although Hal has shown valour in war, he is still best known for his debauched lifestyle and questionable companions. The court is in a state of anxiety over the prospect of his coronation, and Lord Chief Justice is told that he stands 'in coldest expectation' of the new monarch's grace due to his previous run-ins with the prince's oft-unlawful coterie. On one occasion, the play hints, the Chief Justice even sent the prince himself to prison. In spite of this, and going against the advice of Prince Hal's 'sweet' younger brothers, Lord Chief Justice refuses to resort to flattery or to beg forgiveness for his actions. Confronted with the anger of the soon-to-be-crowned Henry V, who has not forgotten his time behind bars, Lord Chief Justice instead asks the nascent king to

Question your royal thoughts, make the case yours;
Be now the father and propose a son
(Henry IV Part 2 Act 5 Scene 2)

Read through the lens of Smith's *Theory*, we can see Lord Chief Justice in the twenty-nine-line passage as appealing to Henry V's capacity for 'sympathy' (TMS I.i.1.5: 13), asking him to bring 'the case home to [his] own breast' (I.i.3.9: 23) by 'changing places in fancy' (I.i.1.3: 12) with his father, Henry IV. Only by so doing, the Chief Justice is saying, can Henry V judge truly whether it was right of Lord Chief Justice to imprison him. He must think whether he himself would have had his own Chief Justice imprison such an insubordinate son:

Hear your own dignity so much profaned,
See your most dreadful laws so loosely slighted,
Behold yourself so by a son disdain'd;
And then imagine me taking your part
And in your power soft silencing your son
(Henry IV Part 2 Act 5 Scene 2)

The spiel works. Henry V finds that he would have had his own insubordinate son imprisoned in such a case, and asks Lord Chief Justice to ‘still carry the balance and the sword’ under his reign. In thus standing up to the king, and in his reply to the ‘sweet princes’, Lord Chief Justice also embodies something of Smith’s ideal ‘wise man’ (TMS III.2.7: 136): A person who so thoroughly identifies with the point of view of the ‘impartial spectator’ (III.2.9: 137) that he cares little for *actual* praise and blame, being instead guided by the desire to be *worthy* of praise—the ‘love of praise-worthiness’ (III.2.25: 147).

In comparison to the other elements of Smith’s theory, the notion of a love of praiseworthiness has received little attention in the secondary literature. This is a significant lacuna. First, without a love of praiseworthiness, the ability to distinguish what is truly praiseworthy from what is merely praised is, morally speaking, worthless. One would know what is right but have no desire to pursue it. Second, Smith’s most substantial discussion of the love of praiseworthiness, added to the sixth and final edition of *TMS*, is mostly concerned with distinguishing this virtuous love from the vain desire for praise; we are not clearly told how we come to have this love, nor how we can foster it in ourselves and in others. Combined, the result is potentially disastrous: Unless we are able to show how the love of praiseworthiness fits within Smith’s wider theory, we are in effect allowing that its plausibility as a normative theory rests on the postulation of an inborn moral compass.

I say ‘potentially’ because, while psychologists up until the late twentieth century were fairly certain that we had to be taught the difference between right and wrong, recent research on the moral psychology of infants actually lend some credence to the notion of an inborn moral compass. Infants as young as six months apparently consider helping others achieve their goals to be good and hindering them to be bad. However, research on infant and child morality also reveals that our early moral compass is rather wonky, its guidance distorted by egotism and in-group partiality. It needs correction through moral education if it is to aspire to anything like a mature morality, not to mention the love of praiseworthiness displayed by Shakespeare’s Lord Chief Justice. If, then, we are not born with a love of praiseworthiness, how does it develop?

In what follows, I will build on insight gleaned from moral psychology, scholarship on Smith, and what Smith himself writes in earlier editions of *TMS* to argue that the development of a full-blown love of praiseworthiness requires the redirection of our naturally strong desire for praise from others toward the imagined impartial spectator, our conscience. The love of praiseworthiness is not an inborn or original love, but a love redirected. Realising this not only helps us make sense of Smith’s theory but aligns it with the current understanding of morality as part natural endowment, part human education. By showing how the love of praiseworthiness relates to the desire for mutual sympathy, my argument also provides a testable hypothesis about moral agency: Developing the ability to experience self-sympathy will correlate positively with individual moral autonomy.

2 Love of praiseworthiness

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. (TMS I.i.1.1: 11)

[Man] desires, not only praise, but praise-worthiness; or to be that thing which, though it should be praised by nobody, is, however, the natural and proper object of praise. (III.2.1: 132)

Both these sentences appear purely descriptive in form. They are statements of fact. In both cases, the appearance of facticity hides normative implications: We might be selfish, but we are not entirely selfish, so when we care for the fate of others, we at least sometimes do so from a genuine concern for them. Likewise, we might be vain, but we are not entirely vain, so we sometimes act with genuine concern for what is right. To the extent that the two statements are similar, it is the second that echoes the first. The

first statement is what meets us when ‘the curtain goes up’ (Griswold 1998: 44) on Smith’s *Theory*. It’s been there since the first edition. The second appears as part of the first paragraph in a chapter that is almost entirely new to the sixth and last edition of *TMS*. In both cases, Smith starts out with a striking claim about moral psychology and proceeds to build upon it a cornerstone of his theory.

In the first case, Smith uses our capacity for sympathy, or ‘fellow-feeling with any passion whatever’ (*TMS* I.i.1.5: 13) to explain our interest in others. Sympathy, in turn, forms the basis for moral judgements since our inability to sympathise with someone equals disapproval of their sentiments as either excessive or deficient (I.i.3.1: 20-21). When Smith opens *TMS* with the claim that we are not wholly selfish, he is confronting the ‘selfish’ systems of morality (VII.ii.4.6-8: 353-55) associated with Thomas Hobbes’ bleak ‘state of nature’ (Hobbes 2009). Smith’s main target is Bernhard Mandeville’s interpretation of this in his *Fable of the Bees* (Mandeville 1962; Frazer 2010:19–22; *TMS* VII.ii.7: 364). Whereas these authors see humans as thoroughly selfish, and any morality as an artificial arrangement of society, Smith is convinced that humans are fundamentally concerned with the wellbeing of others. Smith knows his claim is controversial, and he backs it up with detailed and convincing analysis of the role that sympathy plays in our ability to put ourselves in other peoples’ shoes and to evaluate their actions and reactions on the basis of whether we can or cannot go along with them.

In the second case, Smith builds on his theory of sympathy and the imagined impartial spectator this spawns (*TMS* III: 128-229), and introduces the notion of a ‘love of praise-worthiness’ (III.2.25: 147) to explain how we are able to act in defiance of the opinions of others when our judgements about what ought to be praised differ from theirs. On the one hand, Smith’s second claim – that we not only desire praise but also desire to be worthy of praise even if none is given – can be seen as a corollary of his first: If vanity is a form of selfishness, then the claim that we are not wholly vain is a corollary of the claim that we are not wholly selfish. On the other hand, the claim that we love being praiseworthy carries an explanatory burden not carried by the claim that we are interested in the wellbeing of others. Smith needs this second notion in part because he thinks our first inclination is to seek harmony or concord with those around us. The desire to sympathise, to agree with others, and to have others sympathise with us is what drives the process underlying morality. Our desire for ‘mutual sympathy’ (I.i.2: 17-20) naturally inclines us toward seeking common ground with other people. This explains how the capacity for sympathy can lead to the construction of shared moral norms, and it is, therefore, central to Smith’s analysis of how morality arises from human sentiments.

Morality, however, is not only about agreeing with others, or about following shared norms. A central aspect of moral discourse as we know it is moral disagreement. We are not just trying to get along, but trying to do so in the right way, guided by judgements, arguments and rules that are somehow truer or better than others. This aspect must be accounted for even if the final analysis of what ‘truer’ means refers exclusively to what people happen to feel; we need to understand what it is that allows people to disagree just as much as we need to understand the mechanisms of agreement, regardless of what the agreement or disagreement is about.² While the desire for mutual sympathy carries within it both these tendencies – it is, after all, the compound desire of understanding others and having others understand us – it cannot, on its own, explain how people sometimes completely forego the sympathy of others to pursue what they take to be right. Going against the opinion of people who are close enough to us that we notice their disapproval requires some other motive than just the desire for mutual sympathy with them; it requires an inner strength of some kind that sustains us in our resolve to act on our convictions even when these turn out to be unpopular. When Shakespeare’s Lord Chief Justice stands up to Henry V even though he fears that it may cost him his life, he cannot merely be driven by his belief that he has acted in a manner that is praiseworthy. Coupled only with the desire for mutual sympathy, his belief would fold under the weight of being in discord with the king. What sustains Lord Chief Justice’s resolve faced with the prospect of condemnation from the kingdom’s highest authority is his love of praiseworthiness.

The love of praiseworthiness is in some ways Smith’s answer to the age-old question ‘why be moral?’ – at least if this is understood as a question about why we should care about what is right, or, alternately, as a request for an explanation for why we, in fact, do care. The love of praiseworthiness is our ‘natural

incentive to be virtuous' (Griswold 1998: 130). As such, Smith's claim that we are not wholly vain is controversial in its own right. The duty to account for this is not discharged simply by explaining how it is that we are not wholly selfish. If we are to believe that the second claim is a corollary of the first, we need to know how we get from the one to the other. Smith, unfortunately, does not tell us. In the sixth edition of *TMS*, where Smith writes most about the love of praiseworthiness, he tells us almost nothing about how this love connects to our ability to sympathise, our desire for mutual sympathy, and the imagined impartial spectator.

Connecting what Smith says about the love of praiseworthiness to the other main elements of his theory is, first of all, an exegetical problem. As an exegetical problem, it is compounded by the relative lack of scholarly work on the love of praiseworthiness. While the notions of sympathy and the impartial spectator have been extensively explored in the secondary literature, only a handful of scholars have written explicitly about the nature of this love. One of them is Ryan Patrick Hanley, and although his claim that the love of praiseworthiness is 'logically and temporally prior' to the love of praise is almost certainly false (Hanley 2009: 140), the possibility that the love of praiseworthiness might be an 'original desire' on a par with the love of praiseworthiness deserves further scrutiny (*TMS* III.2.7: 135).

3 An original love?

In his virtue-ethical reading of Smith, *Adam Smith and the Character of Virtue* (2009), Hanley connects Smith's discussion of the distinction between praise and praiseworthiness, and the love of each, to a problem first formulated by Rousseau. Society, thought Rousseau, tends to foster a certain kind of self-love, *amour-propre*, that privileges 'a concern for appearances over a concern for the truth of one's character' (Hanley 2009: 41). The consequence of this is that seeming becomes more important than being: 'in Rousseau's terms, *paraître* supplants *être*' (Hanley 2009: 41). Smith, writes Hanley, saw the same problem arising from certain features of commercial society, features which tend to tie advancement to recognition whether or not that recognition is merited (Hanley 2009: 41).

Smith even makes a distinction quite similar to Rousseau's, namely between *appearing* to be fit for society and actually *being* fit. 'Nature', Smith says, has 'formed man for society', and in so doing has 'endowed him with an original desire to please and an original aversion to offend his brethren' (*TMS* III.2.7: 135). However, 'this desire of the approbation, and this aversion to the disapprobation of his brethren, would not alone have rendered him fit for that society for which he was made'; it 'could only have made him wish to appear to be fit for society' (III.2.8: 136). The original desire to please, which Smith calls the 'love of praise' (III.2.25: 147), will drive people to act in ways that make others see them in a positive light, but only to the extent that there are actual spectators to see and praise them. The act would be just that, an act, a put-on, 'the affectation of virtue, and ... concealment of vice' (III.2.8: 136). The solution to the problem of the separation of *être* from *paraître*, or the explanation for why people are not merely pretending, but genuinely concerned with being virtuous, is the desire to be praiseworthy: 'Nature ... has endowed [man], not only with a desire of being approved of, but with a desire of being what ought to be approved of; or of being what he himself approves of in other men' (III.2.8: 136).

In discussing this part of Smith's argument, Hanley writes that Smith's solution to the Rousseauan problem of the separation of *paraître* from *être* 'rests on his postulation of a love of praiseworthiness *both logically and temporally prior to a love of praise*' (Hanley 2009: 140, my emphasis). To say that our love of praiseworthiness is 'logically and temporally prior' to our love of praise is to say that love of praise depends on the prior existence of a love of praiseworthiness. In other words, if there does not already exist a love of praiseworthiness, neither can there be a love of praise. Does this fit with Smith's theory? It is difficult to see how it could. Indeed, Smith's account of the structure and development of moral agency suggest the opposite ordering. It is our desire to be approved of by others that drives us to view ourselves from without, which in turn lets us see the difference between something being praised and something being worthy of praise. Only after having realised this distinction can we be said to be able to desire to be worthy of praise, and so be able to love praiseworthiness. In this sense at least, it seems that the love of praise is logically and temporally prior to the love of praiseworthiness.³ However, there is a slightly weaker

and more plausible interpretation to be made of Hanley's statement. At the end of the paragraph in which he invokes the logical and temporal priority of love of praiseworthiness, Hanley says that Smith 'makes the Rousseauan claim that nature in its wisdom antecedently invested man with a love of praiseworthiness capable of withstanding and mitigating civilization's corruptions' (Hanley 2009: 140). This, I take it, is a claim that the love of praiseworthiness, like the love of praise, is an 'original desire' (TMS III.2.7: 135), a basic part of human nature not derived from something else, and certainly not derived from the love of praise.

Much of what Smith says about the love of praiseworthiness in the sixth edition of *TMS* can be taken to support this interpretation. For example, after confidently stating that '[man] desires, not only praise, but praise-worthiness' (III.2.1: 132), Smith goes on, in the very next paragraph, to say that '[t]he love of praiseworthiness is by no means derived altogether from the love of praise' (III.2.2: 132). As an example of how the love of praiseworthiness cannot be derived altogether from the love of praise, the third paragraph, also new to the sixth edition, discusses how we judge of our own praiseworthiness by becoming 'the impartial spectators of our own character and conduct' (III.2.3: 133). If what we see from this point of view is 'as we wish', that is, if we can consider ourselves praiseworthy, 'we are happy and contented' (III.2.3: 133). This contentment does not depend on any actual praise from real spectators, but if we were to find that other people see our conduct and character 'in the same light' as we do, '[t]heir praise necessarily strengthens our own sense of our own praise-worthiness' (III.2.3: 133). Smith concludes, '[i]n this case, so far is the love of praise-worthiness from being derived altogether from that of praise; that the love of praise seems, at least in a great measure, to be derived from that of praise-worthiness' (III.2.3: 133). Finally, in the paragraph about appearing versus being fit for society quoted above, Smith says that 'Nature, accordingly, has endowed him, *not only* with a desire of being approved of, but with a desire of being what ought to be approved of' (III.2.8: 136, my emphasis).

The 'original love' interpretation seems to stand on solid ground, exegetically speaking - except, of course, for the fact that Smith never refers to the love of praiseworthiness as 'original', which is an adjective he reserves for our 'original desire to please [our] brethren' (III.2.6: 135). Could this be because the love of praiseworthiness has to be developed from more basic desires? In an endnote to his discussion of the love of praiseworthiness, Hanley points the reader in the direction of Lauren Brubaker's essay *Does the 'wisdom of nature' need help?* (2006) for a fuller treatment of the relationship between nature and the love of praiseworthiness in *TMS* (Hanley 2009: 174, endnote 12). Interestingly, Brubaker's take on the place of the love of praiseworthiness in Smith's theory is almost exactly opposite to that suggested by Hanley. Brubaker says of the impartial spectator and the love of praiseworthiness that they are 'developed or perfected', sentiments that require 'judgement and reflection' to mature (Brubaker 2006: 181). Once perfected, they 'can even come into conflict with the natural sentiments from which they first arose' (Brubaker 2006: 181). Case in point: a properly developed love of praiseworthiness 'can sustain us even under conditions in which we lose the praise of actual spectators, our original or natural desire' (Brubaker 2006: 181).

Brubaker's insight springs from an analysis of Smith's seemingly contradictory uses of the term 'Nature'. Nature, to Smith, is sometimes 'Darwinian' (Brubaker 2006: 177), with its 'favourite ends' being the 'self-preservation and propagation of the species' (TMS II.i.5.10: 90). At other times, Smith imputes to nature a concern for human happiness and perfection (see for example III.5.7: 193) - a concern that sits uneasily with the stark logic of evolutionary adaptation (Brubaker 2006: 177). Brubaker's claim is that this is not a contradiction, but rather the sign of a '*conflict within nature*', a conflict between 'nature simply' and (mature) human nature (Brubaker 2006: 178, emphasis in original; see also Griswold 1998: 313-30). In Smith's own words,

[M]an is by Nature directed to correct, in some measure, that distribution of things which she herself would otherwise have made. The rules which for this purpose she prompts him to follow, are different from those which she herself observes. (TMS III.5.9: 195)

Smith at one point uses the example of an 'industrious knave' and an 'indolent good man' and notes that while nature rewards industry and punishes indolence, it is human nature to prefer the virtuous person to the knave, rendering unjust a natural order in which the knave comes out ahead (III.5.9: 195, see also II.ii.2.1: 97). In this way and others, the rules of nature are liable to 'shock and offend the "natural sentiments" of man' (Brubaker 2006: 178). But nature simply and human nature are not constant enemies. On the contrary, while each follows rules that are fit for them, both sets of rules, Smith tells us, are in fact 'calculated to promote the same great end, the order of the world, and the perfection of human nature' (TMS III.5.9: 196).

Both sets of rules – the laws of nature and the 'law and morality' with which humans intervene in the natural course of things – are needed to promote human happiness for two reasons (Brubaker 2006: 180). First of all, there is no human happiness without humans, and so self-preservation is a precondition for the promotion of human happiness. For our preservation, we depend on 'appetite and instinct' (Brubaker 2006: 180), which are the work of our narrowly Darwinian nature. These internal representatives of nature simply ensure not only our struggle to survive, but also our constant striving to better our condition. The rules of nature are 'useful and proper for rousing the industry and attention of mankind' (TMS III.5.10: 168) since they typically reward industry (Cropsey 1975: 141). Secondly, humans are also naturally social animals, and so the human concern to promote the perfection and happiness of human nature 'through morality, law, and society' is also, in a sense, nature's concern (Brubaker 2006: 181). Immediate appetite and instinct, however, are insufficient to promote human happiness, and '[n]ature simply needs the help of human efforts to correct nature and human nature' (Brubaker 2006: 181). The pursuit or realisation of human happiness 'is "prompted" by nature ... but depends on the wisdom of human efforts' (Brubaker 2006: 181). The impartial spectator and the love of praiseworthiness are core parts of this 'wisdom', so while nature 'sets us on the course to "happiness and perfection" through human society' (or, perhaps: human sociality), achieving that goal means developing standards of judgement that will put us at odds with both nature simply and our own, basic desires, and the motivation to follow these judgements (Brubaker 2006: 181).

A continuation of this conflict between nature simply and human nature can be found in Smith's discussion of the two metaphorical 'tribunals' in which questions of propriety are settled (TMS III.2.32: 150). If we understand the fundamental conflict highlighted by Brubaker as one between egotism and altruism (broadly construed), then the conflict of the two tribunals can be seen as a conflict between the first level of 'law and morality' established by human socialisation and a second level of morality, where the latter is needed to correct for errors in the former. It is a conflict between conventional morality and individual moral autonomy. Smith establishes the moral authority of the first tribunal in Part I of *TMS*, where he shows how, on the basis of sympathy, each of us is made the 'immediate judge' of others (III.2.32: 150). This is the 'inferior tribunal' of 'the man without' (III.2.32: 151, footnote 22; III.2.32: 150), Smith's image for the judgement of society. Smith establishes the moral authority of the second tribunal in Part III, where, on the basis of the distinction between praise and praiseworthiness, he shows how each of us judge ourselves, and in turn each other, from the assumed point of view of an impartial spectator. This is the tribunal of 'the man within' (III.2.32: 150), Smith's image for the judgement of our own conscience. The tribunal of the man within functions as the 'court of appeals' (Freiin von Villiez 2011: 39) for ill-informed or partial judgements by actual spectators. It is therefore only by appeal to this 'much higher tribunal' (TMS III.2.32: 150) that we can achieve some measure of independence from common opinion, some measure of moral autonomy.

In the sixth edition of *TMS*, Smith goes on to say that '[t]he jurisdictions of those two tribunals are founded upon principles which, though in some respects resembling and akin, are, however, in reality different and distinct' (III.2.32: 150). The principles corresponding to the two tribunals are of course the love of praise and the love of praiseworthiness, and Smith is here reiterating his earlier statements about these being 'in many respects, distinct and independent of one another' (III.2.2: 132). In the second edition of *TMS*, however, Smith writes instead that 'if we enquire into the origin of [the tribunal within the breast], its jurisdiction we shall find is in a great measure derived from the authority of that very tribunal,

whose decisions it so often and so justly reverses' (III.2.32: 152, footnote 22). In other words, the love of praiseworthiness is in a great measure derived from the love of praise. The difference between the second and the sixth edition is striking. Does it signal a change of heart on Smith's part? Or is he, more conservatively, emphasising the independence of the love of praiseworthiness as a moral motive? D.D. Raphael has argued that the differences between the first, second and sixth editions are differences in emphasis rather than in the elements of Smith's theory (Raphael 1975: 94, 2007: 44–45).⁴ Taking this view, I will draw on the full spectrum of Smith's varying emphases in reconstructing his argument. The result renders moot the question of Smith's intention in revising: We get a way of seeing the love of praiseworthiness as both derived and independent.

4 A love redirected

To be amiable and to be meritorious; that is, to deserve love and to deserve reward, are the great characters of virtue; and to be odious and punishable, of vice. But all these characters have an immediate reference to the sentiments of others. (TMS III.1.7: 132)

Apart from the identification of virtue with praiseworthiness, the most important thing in this passage is Smith's insistence that virtue has an *immediate* reference to the sentiments of others. He continues, 'Virtue is not said to be amiable, or to be meritorious, because it is the object of its own love, or of its own gratitude; but because it excites those sentiments in other men' (III.1.7: 132). This is probably one of the clearest examples of Smith's non-objectivism about moral value (his skepticism, as Griswold calls it, 1998: 155–73), and it is important to understanding the relationship between love of praise and love of praiseworthiness.

In the paragraphs preceding this, Smith has sketched out the mechanism by which we make judgements about our own conduct and sentiments by analogizing it to how we judge others. We do so, Smith writes, by imaginatively inhabiting the perspective of a spectator: 'We suppose ourselves the spectators of our own behaviour, and endeavour to imagine what effect it would, in this light, produce upon us' (TMS III.1.5: 131). We divide ourselves 'as it were, into two persons ... the examiner and judge [and] the person whose conduct is examined and judged of' (III.1.6: 131). By seeing ourselves from without in this manner, we are able to predict the judgements of others. The analogy is that of a mirror, a 'looking-glass', in which we see how we appear to others, and with the help of which we may adjust that appearance in expectation of their opinion (III.1.4: 130). Thanks to the 'looking-glass' of the imagined spectator, we can congratulate ourselves upon doing something for which we think 'other men' would praise us:

The consciousness that [virtue] is the object of such favourable regards, is the source of that inward tranquillity and self-satisfaction with which it is naturally attended, as the suspicion of the contrary gives occasion to the torments of vice. What so great happiness as to be beloved, and to know that we deserve to be beloved? What so great misery as to be hated, and to know that we deserve to be hated? (II.2.7: 132)

In the first to fifth editions of *TMS*, the paragraph following this, and thus the first paragraph of the chapter entitled 'Of the love of praise, and that of praise-worthiness ...', begins with an assertion: 'The most sincere praise can give little pleasure when it cannot be considered as some sort of proof of praiseworthiness' (III.2.4: 133, see footnote 4: 132). Having explained how the pleasure of being praiseworthy arises from the knowledge that others would praise us if they had the chance, Smith proceeds to discuss how actual praise can give little pleasure if it is not accompanied by an awareness of praiseworthiness.

At first pass, this seems paradoxical: The 'other men' are praising us, and yet we can derive no satisfaction from this because we think that they *would* not. But think of a situation where we are being praised for something we have not in fact done. In that case, our awareness that we have not done anything for which 'other men' would praise us renders the actual praise from those same 'other men' worthless. As Smith writes, '[t]he man who applauds us either for actions which we did not perform, or for motives which had

no sort of influence upon our conduct, applauds not us, but another person' (III.2.4: 133). Here we see clearly how Smith's discussion of praiseworthiness connects to Rousseau's distinction between *être* and *paraître*, as Hanley (2009: 41) points out: To accept unmerited praise is to accept appearing to other people as someone we are not. Letting praise trump praiseworthiness is vanity, plain and simple (TMS III.2.4: 133-34). The obverse of this effect is that the mere knowledge of being praiseworthy 'often gives real comfort ... though no praise should actually be bestowed upon us' (III.2.5: 134), and so we can experience the 'inward tranquillity and self-satisfaction' associated with the knowledge that we have acted virtuously even in the absence of actual spectators.

All the while, the tranquillity offered by our knowledge that we are either praiseworthy or at least not worthy of blame keeps its 'immediate reference to the sentiments of others' (III.1.7: 132). When we do not let ourselves be pleased by unmerited praise, it is because we think that those who praise us *would not* do so had they known what we know. And to the extent that we can endure unmerited blame, it is likewise because we think that all their accusations would fall to the ground *if only they knew*. So, praise- and blameworthiness trump actual praise and blame, not by referring to some objective standard of propriety that exists independently of what anyone might think of it, but by combining what we take to be the actual standard of propriety held by others with the knowledge that we ourselves have of our situation. In other words, if a *well-informed* spectator would praise us, we are praiseworthy. In this manner, our judgements from the point of view of the imagined spectator come not only to predict but to some extent to *supplant* the judgements of others. The imagined spectator thus gains a measure of independence from the actual spectators on which it is modelled.

However, other people will sometimes, even much of the time, fail to praise and blame as they would, not just because they are ill-informed, but because they are *partial*. We are partial to ourselves, and the same is, of course, true of others to themselves. But we are also partial to those close to us, people with whom we identify, certain ideas, books, hairstyles, foods, amongst other things. With these diverse partialities comes an equally diverse set of potential distortions in the way a spectator perceives and judges the conduct or sentiments of a given agent in a particular situation. Thus we often find ourselves in a situation where those 'other men' fail to judge us as they would, not because they lack some crucial bit of information about what we have or have not done (and our motives were), but because their partialities stop them from seeing the situation in the manner that they would, had they not been so partial. And so is added another adjective to the description of the supposed spectator: It must not only be well-informed but also *impartial*.

This, of course, is a double-edged adjective. In the case of being well-informed, there is really no such thing as too much information. The closer we are to the situation and the person judged, the more likely we are to be well-informed. Thus, when predicting what a well-informed spectator would think of our conduct, the fact that we are ourselves, with first-hand knowledge of our situation and our motives, offers a clear advantage. When it comes to being impartial, on the other hand, being ourselves presents more of a problem. Impartiality implies a certain degree of disinterestedness: the ability to judge 'without any prejudice generated either by one's own private emotions of the moment or by any narrow desire to "better one's condition" through manipulation of the situation' (Griswold 1998: 136). That is easier said than done when that 'one' is oneself. Smith has a nice way of putting it when he, in the first edition of *TMS*, discusses the difficulty of regarding oneself with such disinterest:

Unfortunately this moral looking-glass is not always a very good one. Common looking-glasses, it is said, are extremely deceitful, and by the glare which they throw over the face, conceal from the partial eyes of the person many deformities which are obvious to every body besides. But there is not in the world such a smoother of wrinkles as is every man's imagination, with regard to the blemishes of his own character. (TMS III.1.5: 131, footnote 3)

In judging ourselves, the distorting effects of self-love will often cancel out whatever benefit we get from being well-informed about the person we are judging. Nevertheless, the initial independence of the

imagined spectator is at least doubled by the addition of ‘impartial’ to its description. Not only can judgements made from the perspective of a well-informed, impartial spectator supplant the inaccurate judgements of ill-informed, actual spectators, but taking this perspective can also serve as a check both on their *and* on our own self-preferences. Here we see emerging one of the main ‘perfectionist’ elements of Smith’s theory (Forman-Barzilai 2010: 18), namely ‘the man within the breast, the great judge and arbiter of [our] conduct’ (TMS III.2.32: 150). However, the impartial spectator is not, as Forman implies, ‘exogenous’ to Smith’s empirical descriptions (Forman-Barzilai 2010: 18). Rather, it springs forth in the mind of the individual agent as a result of the sympathetic process itself.

When Smith in the sixth edition of *TMS* begins the chapter on the love of praiseworthiness by saying that ‘Man naturally desires, not only to be loved, but to be lovely; or to be that thing which is the natural and proper object of love ... He desires, not only praise, but praise-worthiness’, rather than asserting that ‘The most sincere praise can give little pleasure when it cannot be considered as some sort of proof of praiseworthiness’, (III.2.1: 132), we now know that this is not indicative of the ‘originality’ of the love of praiseworthiness. Instead, Smith is bolstering its claim to independence. With this in mind, the paragraph immediately following can also be seen in a new light:

The love of praise-worthiness is by no means derived altogether from the love of praise. Those two principles, though they resemble one another, though they are connected, and often blended with one another, are yet, in many respects, distinct and independent of one another. (III.2.2: 132)

Considering this a description of the relationship between the two loves after they have entered the scene, Smith’s care to distinguish the love of praiseworthiness from the love of praise makes perfect sense.⁵ The love of praiseworthiness does not simply reduce to a love of praise; the judgements of a well-informed, impartial spectator not only predict, but to some extent supplant the judgements of actual spectators. The imagined spectator represents what real spectators would judge if only they had been *better* spectators. And so emerges not only the independence of the imagined impartial and well-informed spectator but also the normative priority of the judgements of the impartial spectator over those of actual spectators.⁶ To be precise, the reason an impartial judge is normatively superior to a partial judge is that the impartial judge avoids giving preference to any particular agent’s interest in a way that conflicts with the idea that everyone is of equal worth. That idea is inherent in the sympathetic process itself, since by seeing ourselves as a stranger sees us, we come to realise that we are neither more nor less important in their eyes than they in ours: ‘When he views himself in the light in which he is conscious that others will view him, he sees that to them he is but one of the multitude in no respect better than any other in it’ (II.ii.2.1: 97). Our ability to escape the self (Griswold 1998: 78) inexorably puts us on to the truth about our worth relative to that of other people, and, of course, theirs relative to ours.

Put differently, love of praiseworthiness is not derived *altogether* from the love of praise, since if praise comes apart from praiseworthiness, love of praiseworthiness will guide you away from actual praise toward self-approbation through the judgements of the impartial spectator. Love of praiseworthiness is not reducible to the love of praise; our moral motives are not fundamentally selfish. However, love of praiseworthiness *is* derived from love of praise in the sense that it is our desire to be approved of by others that teaches us, by turns, to view ourselves as others see us from the point of view of an imagined spectator, predicting what others *will* judge, then what they *would* judge had they only been well-informed, and, finally, what they *should* judge, as impartial spectators. The love of praiseworthiness is the love of the impartial spectator aspect of ourselves (Griswold 1998: 133). As such, it is a love redirected.

Exegetically speaking, it might be true that the love of praiseworthiness is a love redirected. Smith, however, was an empiricist, and *TMS* is to a large extent an early work of moral psychology, analysing the mechanisms underlying moral judgements. The love of praiseworthiness is no exception to this: It is the name Smith chose for what he took to be one of the driving forces behind the kind of behaviour we usually consider to be praiseworthy or virtuous. The problem of the love of praiseworthiness is therefore also a problem of what empirical support it might have. To go beyond exegesis, and to treat the love of

praiseworthiness as more than a theoretical curiosity, we should, therefore, look to modern moral psychology to test whether the argument that this love is a love redirected fits with our best current understanding of moral development.

5 An inborn moral compass?

In the tradition established by the pioneering work of Jean Piaget and Lawrence Kohlberg, moral psychologists up until the late twentieth century tended to regard moral development as a stadial progression, beginning from an ‘essentially pre-moral’ obedience to authority (Vaish and Tomasello 2013: 279). The question of whether morality might have some innate basis was perhaps not as much dismissed as simply passed over as empirically untestable.⁷ Infants, after all, cannot speak.

The practical hurdle presented by infants’ lack of language was vaulted with the development of experimental paradigms relying on other clues like the amount of time the infants spend looking at different elements in the experiment. The results of this research indicated that even very young infants make moral judgements. One of the pioneering studies of infant morality found that twelve-month-old babies evaluate helping-behaviour as positive, and hindering-behaviour as negative (Premack and Premack 1997: 851–52). More recent studies have found evidence that infants as young as six months (Hamlin, Wynn, and Bloom 2007: 558) and three months (Hamlin, Wynn, and Bloom 2010: 927–30) attribute these behaviours to agents, and that they make corresponding evaluations of the agents as a result: They prefer helpers over hinderers. Reading Karen Wynn’s description of these experiments, it is tempting to label the infants as impartial spectators *in spe*, operating with an innate standard of propriety:

Their evaluations are made on the basis of witnessed interactions between unfamiliar individuals; the infant, as an unaffected, unrelated—and therefore unbiased—third party, is nonetheless rendering an abstract judgment about the value of a social act. (Wynn 2008: 346)

Although the evidence for moral evaluation in very young infants has been called into question (Salvadori et al. 2015; Scarf et al. 2012; see Hamlin, Wynn, and Bloom 2012 for a reply), the idea that we are either born with a rudimentary moral compass (Hamlin 2013) or that we develop one early in childhood (Emde et al. 1991) has support from others sources. There is research showing that toddlers as young as fourteen to eighteen months are ‘naturally altruistic’, and will tend to help others achieve their goals ‘irrespective of any reward from adults’ (Warneken and Tomasello 2009: 455), that infants as young as twelve months prefer equal to unequal distribution of goods among third parties and equal over unequal distributors (Geraci and Surian 2011: 1016–17), and that in fifteen-month-olds, a sensitivity to fair distribution is correlated with altruistic behaviour (Schmidt and Sommerville 2011: 5). Twenty-one-month-old babies even take into account relative merit in their evaluations of fair distributions (Sloane, Baillargeon, and Premack 2012: 203). Children of ages four to seven are also averse to getting less than others in a distribution of goods (Blake and McAuliffe 2011: 211), and when children three to five years old are themselves responsible for the distribution of goods, they do in fact distribute, that is, they do not simply take everything for themselves even if they could (Rochat et al. 2009: 427).

Could the love of praiseworthiness be an original or innate love after all? Other experiments complicate the picture. Wynn (2009) reviews a number of studies on how infants prefer individuals who are like themselves over those who are dissimilar and argues that infants are probably not as indiscriminately altruistic as Warneken and Tomasello (2009) suppose. Hamlin *et al.* (2013) followed up on this by introducing an element of social identification in their experimental setup: Prior to being exposed to a scenario in which a puppet is helped or hindered, the infant was shown that the puppet either shared or did not share the infant’s taste in snacks. If the infant preferred green beans and the puppet preferred graham crackers, both nine- and fourteen-month-old infants would subsequently want the puppet harmed rather than helped (Hamlin et al. 2013: 592). It is not just that infants who love beans are evil; the desire to see the dissimilar other harmed was just as strong for those who chose the crackers. More importantly, the infant’s desire to see the other *helped* was just as strong as it had been in the original experiment if the

puppet shared the infant's taste in snacks (Hamlin et al. 2013: 590). This result is hard to square with the idea of an original love of praiseworthiness 'capable of withstanding and mitigating civilization's corruptions' (Hanley 2009: 140), but it is less than surprising considering the factors in the evolutionary history of humans which are likely responsible for shaping any innate tendencies humans might have, like kin (or even group) selection (Joyce 2006: 45–47; Lewontin 1970: 12–16; Vaish and Tomasello 2013: 280–82). In contrast, the interpretation that the love of praiseworthiness is a love redirected, especially if this is seen through the lens of the 'conflict within nature' that Brubaker identifies in Smith's writing, is altogether compatible with the finding that infants' innate tendencies are only imperfectly moral. If there is anything like a love of praiseworthiness in us from birth, it is very much the poor relation of the love of praiseworthiness described by Smith.

The rosy image of the fair and altruistic infant is further marred by a second set of results from studies on distributive fairness. At the same age where children object to the inequality of getting less than others, they are not averse to getting significantly more (Blake and McAuliffe 2011: 211; Rochat et al. 2009: 441). Moreover, five- to six-year-old children will choose to receive fewer goods in absolute terms if they thereby ensure that they get relatively more than an anonymous peer (Sheskin, Bloom, and Wynn 2014). Sheskin *et al.* suggest that the observed 'anti-equality' in young children might be attributed to the influence of social comparison: children are not concerned to get as much as possible, but to get more than others, even if that means getting less than they otherwise would (2014: 155). Babies, it turns out, are not impartial spectators of themselves.

As for when the tendency to impartiality develops, studies of distributive fairness indicate that a change happens around the eight-year mark. Children of this age will prefer an egalitarian over unequal distribution of goods (Fehr, Bernhard, and Rockenbach 2008: 1082), and reject a distribution of candies where they themselves get significantly more than an anonymous other (Blake and McAuliffe 2011: 211). Children also gradually develop sensitivity to the importance of impartiality in procedures (Shaw and Olson 2014: 48) and an understanding of its importance in judges (Mills and Keil 2008: 544–47). This further supports the conclusion that it takes time to develop the habit of impartial spectatorship.

Finally, and suggestively for the specific question of the development of a love of praiseworthiness, a pair of studies have looked at the disconnect between the recognition that a norm applies equally to others and to oneself and the desire or willingness to follow it. The first study found that children five to six years old are aware of fairness norms, but prefer to act selfishly; children seven to eight years old act fairly, but derive no pleasure therefrom; and children nine to ten years old both act fairly *and* derive pleasure from equal distribution (Kogut 2012: 235–36). The second study found that children three to eight years old recognise that fairness norms applied equally to themselves and others, but that the younger children failed to follow these norms in practice (Smith, Blake, and Harris 2013: 7). The failure was not due to weakness of will in the moment. Asked to predict whether they would follow fairness norms in future distributions, children under seven correctly predicted that they would *not*. Children seven to eight years old both predicted that they would follow fairness norms and actually did follow them in practice (Smith et al. 2013: 8). The authors hypothesise that the developmental change underlying this effect is not an increased ability to inhibit the impulse to self-satisfaction, but rather an increased regard to the weight of normative considerations in situations where there is tension between norms and selfish desires (Smith et al. 2013: 8). In Smithian language, children only gradually develop a love of praiseworthiness that is sufficiently strong to overrule their original self-love.

In sum, even very young infants make judgements about praiseworthiness that appear to be independent of any moral education. However, the moral compass of early life puts them off course through the unchecked influence of self-love and love for similar others. It takes time and effort – about three to five years' worth – to get into the habit of viewing oneself from the perspective of the impartial spectator, and it takes even longer – about eight years – to develop a robust desire to act on what one then sees. Charles Griswold's gloss on Smith's central argument neatly captures the extent to which Smith's eighteenth-century brand of sentimentalism joins up to twenty-first-century moral psychology:

Smith's argument is ... that the fundamental structures and psychology of moral judgment provide us with a means of distinguishing between [praise- and blameworthiness] and also with a natural inclination to do so. It is an inclination to be realized through moral education, such that the impartial spectator's practical reason becomes our own, becomes (as it were) our second nature. (Griswold 1998: 131)

The empirical evidence supports the exegetical conclusion: Love of praiseworthiness is not an inborn, original love, but must be developed from more basic inclinations. How, then, can we develop it?

6 Self-sympathy

The key to providing a Smithian answer to this question lies in Smith's notion of mutual sympathy. Given what I said in section 2 about how the love of praiseworthiness acts as a counterbalance to the desire for harmonious relations with others, this might seem paradoxical. However, the love of praiseworthiness is itself a version of the desire for mutual sympathy.

To see this, consider how approving of your own conduct relates to sympathy. Your ability to judge of the propriety of somebody else's sentiments or conduct is based on your ability or inability to sympathise with them, to go along with them in what they are feeling and doing. Your ability to judge yourself is similarly based on your ability to sympathise with your own sentiments and conduct from the imagined point of view of an impartial spectator. If you, in viewing yourself from this point of view, find yourself to be praiseworthy, you are therefore sympathising with yourself. You are, in effect, in a state of mutual sympathy with yourself, or, more precisely, in a state of mutual sympathy between the agent and spectator aspects of yourself. As Griswold writes,

The love of virtue is not the love of the approval of some other person, called the 'impartial spectator', but of an aspect of ourselves with which we 'sympathize.' At this level, it is a question of the self's relation to itself ... The love of virtue is an outgrowth of sympathy. (Griswold 1998: 133)

The love of praiseworthiness can, therefore, be seen as the desire for mutual sympathy between these two sides. The love of praiseworthiness is the desire for a state in which you, as an impartial spectator, are able to go along with yourself, as an agent, in what you are doing, feeling and thinking. It is, in short, the desire to do what you believe is right.

In her reconstruction of Smith's argument, Carola Freiin von Villiez (2006, see also 2011) identifies the desire for sympathy as the main motivating principle in *TMS*. She divides the desire for sympathy into three different levels, corresponding to three different stages of the development of moral agency. First, there is the 'instinctive' sympathy of emotional contagion and 'affective communication' with other people, where our 'natural desire' is to be pleasing to those around us (Freiin von Villiez 2006: 149–50).⁸ Second, there is the desire for sympathy with a 'virtual spectator', an internalised representative of conventional morality (Freiin von Villiez 2006: 150–51). Third, there is the desire for sympathy with the 'ideal spectator', the mode of moral judgement in which we critically reflect on the conventional norms of society, and possibly reject them as unjustified (Freiin von Villiez 2006: 152). The first and second levels of sympathy can be associated with the love of praise, and the second and third levels with the love of praiseworthiness. The reason why the distinction between the two loves does not neatly map onto the three levels identified by Freiin von Villiez is that a properly developed 'virtual spectator' – the target of the desire for sympathy at the second level – is an approximation to an imagined impartial spectator, and thus a source of judgements of praiseworthiness (Freiin von Villiez 2006: 151). What changes in terms of motivation when we move from the second to the third level is the renunciation of 'collectively justified approval' (Freiin von Villiez 2006: 151), or the expectation of actual, merited praise. In the words of my argument, acting purely from a desire for mutual sympathy with an ideal impartial spectator means acting purely from a love of praiseworthiness.

The love of praise and the love of praiseworthiness are thus really only names for different modes of the desire for mutual sympathy. What changes between them is the spectator whose sympathy you desire: other people or yourself, judging under the perfectionist aspiration to be an impartial spectator. This is also the most precise sense in which love of praiseworthiness is a love redirected: Moral maturation consists, at least in part, in redirecting your desire for mutual sympathy away from other people toward your own conscience.

Fostering a love of praiseworthiness thus means fostering a facility with self-sympathy. Having a facility with self-sympathy should, therefore, be positively correlated with moral autonomy. This is a testable hypothesis. Given the extent to which Smith's *Theory* accords with contemporary moral psychology, there is, I think, good reason to put it to the test.

7 Conclusion

No one is born with the moral character of a Lord Chief Justice. The love of praiseworthiness is not antecedently invested in us by wise Nature, but must be cultivated from more basic natural inclinations. Fortunately, these natural inclinations are neither wholly selfish, nor wholly vain.

Developing the ability to view oneself impartially takes time, and the desire to act in accordance with what we then see even longer. Indeed, most of us will probably not develop this desire sufficiently to never be consciously and unjustifiably partial. Moreover, the love of praiseworthiness has its dark side. Insensitivity to actual praise and blame, which we could call moral arrogance, can lead us astray just as surely as can vanity, and with greater potential for dire consequences. Smith gives clear-headed analyses of this all-too-common corruption of our moral sentiments in his discussions of false religion (TMS III.5.12: 205-7) and fanaticism (III.3.43: 180-81).

And yet, the two great aims of moral maturation must be the ability to see situations of moral import in the right way and to desire to act on what one then sees – to be appropriately impartial and to love praiseworthiness. According to Smithian sentimentalism, the first aim can only be achieved by continually interrogating the moral assumptions of yourself and those around you, always striving to understand the situation at hand and view it from an impartial point of view. It is to the second aim, without which the first is of little practical value, that this article speaks. The message is that we can only hope to achieve this aim if we leverage our desire for approval from others, combining it with the voice of our conscience so that we find satisfaction in our ability to go along with ourselves in our sentiments and conduct. It is not by recognising one's grim duty, but by learning to take pleasure in doing the right thing that one will develop the moral character of a Lord Chief Justice.

That, at least, is a testable hypothesis.

Bibliography

- Blake, Peter R. and Katherine McAuliffe (2011) “‘I Had so Much It Didn’t Seem Fair’”: Eight-Year-Olds Reject Two Forms of Inequity’, *Cognition*, 120(2): 215–224.
- Brubaker, Lauren (2006) ‘Does the “wisdom of Nature” Need Help?’, in Leonidas Montes and Eric Schliesser (eds.), *New Voices on Adam Smith, Routledge Studies in the History of Economics*, London: Routledge Press, pp. 168–92.
- Cropsey, Joseph (1975) ‘Adam Smith and Political Philosophy’, in Andrew S. Skinner and Thomas Wilson (eds.), *Essays on Adam Smith*, Oxford: Clarendon Press, pp. 132-53.
- Emde, Robert N., Zeynep Biringen, Robert B. Clyman, and David Oppenheim (1991) ‘The Moral Self of Infancy: Affective Core and Procedural Knowledge’, *Developmental Review*, 11(3): 251–70.
- Fehr, Ernst, Helen Bernhard, and Bettina Rockenbach (2008) ‘Egalitarianism in Young Children’, *Nature*, 454(7208): 1079–83.
- Forman-Barzilai, Fonna (2010) *Adam Smith and the Circles of Sympathy: Cosmopolitanism and Moral Theory*, Cambridge: Cambridge University Press.
- Frazer, Michael L. (2010) *The Enlightenment of Sympathy: Justice and the Moral Sentiments in the Eighteenth Century and Today*, Kindle edition, New York: Oxford University Press.
- Freiin von Villiez, Carola (2006) *Dimensionen Der Unparteilichkeit: Adam Smith Auf Der Suche Nach Dem Moralischen Standpunkt*, Habilitationsschrift, Universität Bremen.
- Freiin von Villiez, Carola (2011) ‘Adam Smith’s Story of Moral Progress’, *The Adam Smith Review*, 6: 30–45.
- Geraci, Alessandra and Luca Surian (2011) ‘The Developmental Roots of Fairness: Infants’ Reactions to Equal and Unequal Distributions of Resources’, *Developmental Science*, 14(5): 1012–20.
- Griswold, Charles L. (1998) *Adam Smith and the Virtues of Enlightenment*, Kindle Edition, Cambridge: Cambridge University Press.
- Hamlin, J.Kiley (2013) ‘Moral Judgment and Action in Preverbal Infants and Toddlers Evidence for an Innate Moral Core’, *Current Directions in Psychological Science*, 22(3): 186–93.
- Hamlin, J.Kiley, Neha Mahajan, Zoe Liberman, and Karen Wynn (2013) ‘Not Like Me = Bad Infants Prefer Those Who Harm Dissimilar Others’, *Psychological Science*, 24(4): 589–94.
- Hamlin, J.Kiley, Karen Wynn, and Paul Bloom (2007) ‘Social Evaluation by Preverbal Infants’, *Nature*, 450(7169): 557–59.
- Hamlin, J.Kiley, Karen Wynn, and Paul Bloom (2010) ‘3-Month-Olds Show a Negativity Bias in Their Social Evaluations’, *Developmental Science*, 13(6): 923–29.
- Hamlin, J.Kiley, Karen Wynn, and Paul Bloom (2012) ‘Reply to Scarf et Al.: Nuanced Social Evaluation: Association Doesn’t Compute’, *Proceedings of the National Academy of Sciences*, 109(22): E1427–E1427.
- Hanley, Ryan Patrick (2009) *Adam Smith and the Character of Virtue*, Kindle Edition. Cambridge: Cambridge University Press.

- Hobbes, Thomas and C.B. MacPherson (Introduction) [1651] (1986) *Leviathan: The Matter, Forme, & Power of a Common-Wealth Ecclesiastical and Civill*, London: Penguin Books.
- Joyce, Richard (2006) *The Evolution of Morality*, Cambridge, MA.: MIT Press.
- Kogut, Tehila (2012) 'Knowing What I Should, Doing What I Want: From Selfishness to Inequity Aversion in Young Children's Sharing Behavior', *Journal of Economic Psychology*, 33(1): 226–36.
- Lewontin, Richard C. (1970) 'The Units of Selection', *Annual Review of Ecology and Systematics*, 1: 1–18.
- Mandeville, Bernard [1714] (1962) *The Fable of the Bees: Private Vices, Publick Benefits*, ed. Irwin Primer, New York: Capricorn Books/G. Putnam's Sons.
- Mills, Candice M. and Frank C. Keil (2008) 'Children's Developing Notions of (Im)partiality', *Cognition*, 107(2): 528–51.
- Premack, David and Ann James Premack (1997) 'Infants Attribute Value± to the Goal-Directed Actions of Self-Propelled Objects', *Journal of Cognitive Neuroscience*, 9(6): 848–56.
- Raphael, D. D. (1975) 'The Impartial Spectator', in Andrew S. Skinner and Thomas Wilson (eds.), *Essays on Adam Smith*, Oxford: Clarendon press, pp. 83–99.
- Raphael, D. D. (2007) *The Impartial Spectator*, Oxford University Press.
- Rochat, Philippe et al. (2009) 'Fairness in Distributive Justice by 3-and 5-Year-Olds across Seven Cultures', *Journal of Cross-Cultural Psychology*, 40(3): 416–442.
- Scarf, Damian, Kana Imuta, Michael Colombo, and Harlene Hayne (2012) 'Social Evaluation or Simple Association? Simple Associations May Explain Moral Reasoning in Infants', *PLOS ONE*, 7(8): e42698. doi:10.1371/journal.pone.0042698.
- Schmidt, Marco F. H. and Jessica A. Sommerville (2011) 'Fairness Expectations and Altruistic Sharing in 15-Month-Old Human Infants', *PLoS ONE*, 6(10): e23223. doi: 10.1371/journal.pone.0023223.
- Shaw, Alex and Kristina Olson (2014) 'Fairness as Partiality Aversion: The Development of Procedural Justice', *Journal of Experimental Child Psychology*, 119: 40–53.
- Sheskin, Mark, Paul Bloom, and Karen Wynn (2014) 'Anti-Equality: Social Comparison in Young Children', *Cognition*, 130(2): 152–56.
- Sloane, Stephanie, Renée Baillargeon, and David Premack (2012) 'Do Infants Have a Sense of Fairness?', *Psychological Science*, 23(2): 196–204.
- Smith, Adam [1751/1790] (2002) *The Theory of Moral Sentiments*, ed. K. Haakonssen. Cambridge: Cambridge University Press.
- Smith, Craig E., Peter R. Blake, and Paul L. Harris (2013) 'I Should but I Won't: Why Young Children Endorse Norms of Fair Sharing but Do Not Follow Them', *PLOS ONE*, 8(3): e59510. doi:10.1371/journal.pone.0059510.
- Vaish, Amrisha and Michael Tomasello (2013) 'The Early Ontogeny of Human Cooperation and Morality', in Melanie Killen and Julia G. Smetana (eds.), *Handbook of Moral Development*, New York and London: Psychology Press, pp. 279–98

Warneken, Felix and Michael Tomasello (2009) 'The Roots of Human Altruism', *British Journal of Psychology*, 100(3): 455–71.

Wynn, Karen (2008) 'Some Innate Foundations of Social and Moral Cognition', in Peter Carruthers, Stephen Laurence, and Stephen Stich (eds.), *The Innate Mind, Volume 3: Foundations and the Future*, Oxford: Oxford University Press, pp. 330–347.

Wynn, Karen (2009) 'Constraints on Natural Altruism', *British Journal of Psychology*, 100(3): 481–85.

1Notes

.All references to *TMS* will be to the Cambridge edition (2002).

- 2 .Even if, as an emotivist or error theorist would claim, the disagreement is not *about* anything, we would still like to know, for example what people are thinking when they are disagreeing. The psychology is interesting independently of the metaphysics.
- 3 .At the *Smith and Virtue* workshop in Oslo, august 2015, Hanley responded to a version of this criticism by saying that, had he written the paragraph again, he would have said that love of praiseworthiness is *lexically* prior to the love of praise. Lexical priority is common way of ranking principles in ethics, especially in the tradition after John Rawls. If we interpret Hanley's claim in this light, we could say that love of praiseworthiness is morally more important than love of praise. If in a particular situation there is a conflict between acting so as to be worthy of praise and so as to be praised, one ought to act from a love of praiseworthiness rather than a love of praise. This, I think, is a wholly unproblematic claim about Smith's model of moral judgement. However, it fails to answer the question of why or how it is the case that humans not only love praise, but also love praiseworthiness, and, even more in need of an explanation: how we come to hold praiseworthiness to be more important than praise.
- 4 .Note that Raphael is primarily talking about the status of the impartial spectator and how Smith's descriptions of this changes throughout the different editions. Given the intimate connection between the impartial spectator and the love of praiseworthiness, to be explored later in this article, I think the same argument can be extended to Smith's changing descriptions of this love.
- 5.Compare also Smith's use of "altogether" here and in VI.ii.2.4: 270.
- 6.Or, as I assume Hanley would phrase it: the lexical priority.
- 7 .As Vaish and Tomasello point out, Piaget was careful to note that he was studying explicit moral judgement, not moral behaviour or sentiments – the more recent research thus complements rather than competes with this tradition (2013: 280).
8. All translations from the German original by the author.