'That a feeling is bestowed on us by Nature, does not necessarily legitimate all its promptings. The feeling of justice might be a peculiar instinct, and might yet require, like our other instincts, to be controlled and enlightened by a higher reason. If we have intellectual instincts, leading us to judge in a particular way, as well as animal instincts that prompt us to act in a particular way, there is no necessity that the former should be more infallible in their sphere than the latter in theirs: it may as well happen that wrong judgments are occasionally suggested by those, as wrong actions by these.' Discuss.

1. In this opening to *Utilitarianism*'s chapter on justice, Mill argues that instincts do not necessarily lead us to right actions – and consequently, that reasoned thought plays an essential role in developing our moral theories. Mill's stance against pure intuitionism, the view that the truth-values of normative sentences are self-evident and accessible via a special mental faculty, seems justified, not least because it is unclear in what sense any of these truth-values would be self-evident. In this essay, I shall first briefly explain the substance of Mill's argument in the extract, before laying out in greater detail why I believe his claims are reasonable, and finally exploring how this piece of reasoning fits in with the broader theses of *Utilitarianism.*

2. Mill's argument in this extract can be summarised as follows:

> **P1:** Our instincts do not always lead to correct conclusions or right actions.
>
> **P2:** Feelings of justice are instinctive.
>
> **P3:** To reach correct conclusions and right actions we must apply practical reason to our instincts.
>
> *(**A:** We have an obligation to reach correct conclusions and right actions.)*
>
> **C:** We must apply practical reason to our feelings of justice.

A is an assumption introduced by Mill earlier on in *Utilitarianism* (2.10), while the other premisses are spelled out explicitly. I will not consider objections to **A** in the following, since they are not directly relevant to our examination of intuitionism; someone opposing **A** would have more fundamental disagreements with Mill's work than the arguments raised in the excerpt above.

3. In order to properly assess the soundness of **P1,** and inform our forthcoming discussion of intuitionism, it is worth spelling out the semantic difference between "instinct" and "intuition": the former tends to refer to an impulsive, perhaps biologically-preprogrammed, response to a stimulus, whereas the latter typically denotes less hasty and more considered (though still subconscious) cognitive processing. Mill's argument here is framed around instincts, which makes defence of **P1** simple, since it is quite straightforward to put forward examples of when an instinct (say, to retaliate in a pub brawl) leads to wrong actions. The same can be said for moral instincts, as Nozick (p162) demonstrates neatly in his thought experiment concerning whether citizens should be permitted to work overtime and accrue additional resources: whilst the resulting inequality may feel instinctively unfair, upon further consideration it might seem more like a just state of affairs. The question of whether our *intuitions* can be relied upon will return when we analyse **P3**, but for now it is enough to conclude that **P1**, in any case, seems justified and sound.

4. Mill's convincing justification of **P2** is done through a quasi-anthropological explanation of the origins of our instinct for justice. The motivation here is that by establishing a plausible natural cause for the sentiment, Mill can dispel the notion that it is a special source of objective moral information which may be in conflict with his utilitarianism (Crisp p156). Mill suggests that the relevant natural cause is our animal desire for retribution and vengeance (5.21). It is tempting to think that Rawls (amongst others) would object to this characterisation and thus **P2**, arguing instead that justice is the product of our mental capacities (p42). But one must be careful not to muddle the two concepts being discussed:

when Mill talks about the fallibility of our "sentiments of justice", he doesn't mean to say that justice leads us to wrong actions – in fact, Mill is explicit that true justice is "always virtuous" (5.37). Rather, Mill is arguing that instinctive moral judgements made on the basis of what we fleetingly *feel* to be just may often be mistaken. It is the very fact that, as **P2** compellingly asserts, our feelings of justice are instinctive which mean that mental capacities are required to determine what truly is just. Mill's naturalistic analysis, leading us to conclude that our instincts of justice do not contain objective moral truths, can sit comfortably alongside Rawls' conception of justice.

5. In fact, there are significant similarities between Mill's views about the proper way to produce an ethical theory (as expressed in **P3**) and Rawls's – as well as good reason to support both. Rawls's technique of reflective equilibrium (pp18-19) is an apposite example of practical reason being applied in conjunction with our moral instincts to derive coherent general principles: the thinker starts with a potential set of principles, reasons them through to their practical implications, compares that with their instinctive judgements, and then iteratively adjusts the either the principles or their judgements to iron out any discrepancies. This process of teasing out fundamental axioms from our instinctive judgements via reasoned introspection is precisely what Mill endorses in **P3**, and it is difficult to see how else we could access our deeply-held moral principles other than through such a procedure.

6. We have shown that Mill's argument is sound, and that our instinctive feelings of justice must be subject to practical reason. What part, though, does this play in Mill's broader mission in *Utilitarianism*? Mill succeeds in demonstrating that our sentiments of justice can sit within a utilitarian framework, rather than acting as an additional source of moral information, and that so doing avoids introducing undesirable complications around how one balances competing intrinsic goods of welfare and justice (Rawls p33). This matters because Mill argues in 1.5 that ultimate ends cannot be proved directly, so his justification for utilitarianism at least partly relies on the fact that its (singleton) set of ultimate ends (utility) is simpler than any others, and yet (he hopes to have shown) sufficiently morally expressive. We can draw a parallel with how Smart (pp4-5) does not attempt to prove act-utilitarianism, but settles instead with illustrating the complexities and deficiencies in its alternatives. It therefore seems clear that the ultimate purpose of this extract, for Mill, is more about beating down competing ethical theories than promoting utilitarianism directly.

7. To conclude, the extract demonstrates Mill's strong opposition to the idea of a moral sense, and his support for the central role played by practical reason in reaching moral judgements. Although other philosophers, notably Rawls and Nozick, formed very different conceptions of justice to Mill, a common thread is their use of introspection and reasoning to reach those conceptions. Mill's metaethical arguments in the passage are therefore convincing, for the reasons outlined above.

References

Roger Crisp, "Routledge Philosophy Guidebook to Mill on Utilitarianism" (Routledge, 1997): 95-124.

J.S. Mill, "Utilitarianism" in "The Blackwell Guide Mill's Utilitarianism" ed. Henry R. West (Blackwell, 2006).

Robert Nozick, "Anarchy, State, and Utopia" (Blackwell, 1977).

John Rawls, "A Theory of Justice" (Harvard, 1999 ed.).

J.J.C. Smart, "An Outline of a System of Utilitarian Ethics" in Smart & Williams, "Utilitarianism: For and Against" (Cambridge, 1993).