

Blair's criticism: Nichols owes us a deeper account of his 'normative theory'.

"How precise do the specified conditions have to be? Is it ever acceptable to harm a child? Is it acceptable to harm a child to improve her table manners? Or is it only acceptable to harm a child to prevent her engaging in life-threatening activities? Moreover, it remains unclear how this normative theory develops. Should there be individual differences in this normative theory? If there should be, why should there be?" (pg 276)

Blair claims that no such theory is required for a sentimentalist approach. For Blair, all that is required is to consider the interaction of the neural systems involved in the emotional response to the transgression situation with those involved in the theory of mind (where 'theory of mind' refers to the ability to represent the mental states of others). In other words, the idea is that a moral judgment could arise by one part of the brain having an affective response, and another part of the very same brain representing the mental state of another.

Wrong action: When the action was intended to cause harm.

Bad Action: When the action was not intended to cause harm.

Blair	Nichols
Step 1: Assess ¹ victim distress (affective)	Step 1: Assess victim distress (affective)
Step 2: Assess transgressor's intention (theory of mind)	Step 2: Check normative theory to see if what caused the distress is proscribed (a body of mental representations that has survived cultural evolution)
Step 3: Make moral judgment	Step 3: Make moral judgment

As a final bit of pressure on Nichols, Blair notes that while we certainly can verbalize certain normative rules, that does not make it clear that those rules play a role in how we reason about moral issues. Hence, though such a normative theory such as Nichols proposes might exist, it is not necessarily part of how we arrive at moral judgments.

¹ My term.

D'arms Criticism: Nichols was wrong to reject neo-sentimentalism.

Nichols' theory is what D'arms calls a sentimentalist rules theory (SRT). Specifically, Nichols argues (1) that there are empirical grounds for thinking that neo-sentimentalism is inconsistent with attributing the capacity for moral judgment to young children, (2) that this is a significant cost of the theory, (3) that his own preferred theory, the sentimental rules account, avoids this difficulty, and (4) that his theory meets the list of desiderata that motivate the move to neo-sentimentalism in the first place. D'arms will grant (1) and (3) for purposes of this discussion and focus his critical remarks on (2) and (4).

Response to (2)

Neo-sentimentalism: a family of theories holding that various evaluative or normative judgments are judgments about the appropriateness of sentimental responses (specifically guilt and anger)

For Nichols, the *dissociation problem* was a decisive blow to the neo-sentimentalist project. "If moral judgments are judgments of the appropriateness of guilt, then *an individual cannot have the capacity to make moral judgments unless she also has the capacity to make judgments about the appropriateness of guilt*" (pg 259). According to the suggestion of many studies, however, children ages 3-4 can make moral judgments without understanding complex emotions such as guilt. Hence, moral judgments can be made without understanding the appropriateness of guilt. Hence, neo-sentimentalism is wrong.

D'arm's response to this is quite similar to Smith's response to the amoralist. Children may use the terms 'right' and 'wrong' in an extensionally adequate way (i.e. they appear to have mastery of moral terms), but that alone is not sufficient evidence that they are making moral judgments. Evidence of this can be seen when children use terms that they have picked up from family / friends / other sources despite not having any idea what the term means. Hence, the studies which suggest what Nichols claimed are not really evidence of anything, and the dissociation problem is not decisive.

Response to (4)

The benefits that Nichols claims for SRT are also mostly, if not entirely, attributable to neo-sentimentalism.

- Gibbard, for example, could happily accept that participants in moral discourse do have some shared body of norms, which may just be equivalent to Nichols' body of internally represented rules.
- Both SRT and neo-sentimentalism must suppose that we have some way of recognizing which harms are prohibited and of coding those prohibitions that we regard as moral. And of course, the neo-sentimentalist will take it that the latter coding reflects one's attitudes toward feeling guilt and anger over the action—indeed, he will take it that this connection is what makes it a moral coding at all (whereas I am not sure what

distinguishes moral prohibitions from other prohibitions on the sentimental rules account—which D'Arms take to be an important problem.²)

- Nichols' claim that neo-sentimentalism is 'spectacularly intellectualized' is underdeveloped. (i.e. besides the case of children aged 3-4, Nichols has no other basis for this claim)
- Additionally, some kind of dispositional account of moral reasoning can handle worries about the account being overly intellectual while retaining the neo-sentimentalist thesis

Disagreement

D'Arms claims that SRT cannot account for moral disagreement in the way that Nichols claimed. Nichols cited Stevenson's 1937 but only mentioned one of his two ways of thinking about moral disagreement. Stevenson was targeting *community subjectivism* (CS) whereby moral judgments just are reducible to what the community has rendered permissible / impermissible. Nichols noted that, according to Stevenson, if moral judgments are just judgments about what a community disapproves of, then disagreements between two members of *different* communities is senseless. Since such disagreement is not senseless, CS is false. The second way that Stevenson raises the problem of moral disagreement, which Nichols does not mention, is in accounting for intra-communal disagreement (disagreement between members of the same community). This second problem would seem to be resolvable by just taking a poll, and this is a silly way to think that moral disagreement should be resolved. Since this is a silly way to think that intra-communal disagreement could be solved, CS is false.

D'Arms focuses on the scope of Nichols' normative theory. If such a theory is just confined to a particular community, then Stevenson's first problem immediately arises. If disagreement occurs within a group as to which rules are the correct ones, then we face Stevenson's second problem. D'Arms, however, claims that Gibbard's theory can avoid these challenges.

On Gibbard's analysis, this comes down to a disagreement in attitude: Their attitudes are in conflict because they endorse different norms for when guilt and anger are warranted. The idea that what makes moral disputes univocal, despite dramatic differences in the substantive convictions of the parties, is that they concern the appropriateness of responses to which all parties to the dispute are susceptible. In other words, a moral disagreement is a disagreement about if we are having the right attitude towards an action. Inter-communal disagreement, then, is disagreement about which norms and attitudes are correct, and the same would hold for intra-communal disagreement.³

Objective?

Interestingly, D'Arms notes that Nichols (in his book) does not think that there is a way to cash out his SRT in objective terms. Contrast that characterization with the following quote from the end of the Nichols that we read:

² Perhaps Nichols could appeal to his 'fitness' for 'survival' through 'cultural evolution'?

³ I'm not entirely sure how this avoids Stevenson's second problem. Thoughts?

“As with “basic emotions” like sadness, anger, disgust, and fear, there is good reason to suppose that the emotional response to suffering in others is universal and innately specified. As a result, we should expect that in all cultures, harming people will tend to produce seriously aversive affect. Thus, harmful actions themselves will be likely to arouse negative affect, all else being equal. Just as we’ve seen that norms prohibiting disgusting actions have been extremely successful, so too have harm norms done well historically. It has become a commonplace in discussions of moral evolution that, in the long run, moral norms exhibit a characteristic pattern of development. First, harm norms tend to evolve from being restricted to a small group of individuals to encompassing an increasingly larger group. That is, the moral community expands.” (pg 272)

Does this sound like the groundwork for an objective SRT?