

The "Invariant Sequence" Hypothesis:  
Ethics as an Experimental Science

Todd R. Davies  
Psychology 211  
March 20, 1988

In this paper I will be considering some arguments made by Lawrence Kohlberg in two papers (Kohlberg, 1971 "[A]"; Kohlberg, 1976 "[B]") concerning the existence of an invariant developmental progression ("invariant sequence"--Kohlberg's term, [B] p. 42) in moral reasoning, methods for validating its existence, and its support for an objectivist ethics. A good deal of the debate surrounding Kohlberg's position has focused on his claim that if one can establish empirically that moral reasoning in humans always (for "non-philosopher subjects" [A]) develops along one progression or pattern, then the principles that are held in the last stage (which not all subjects must attain) must be the most moral. If Kohlberg is right about this, then there are serious consequences for the philosophical enterprise of ethics, namely that if such an invariant sequence can be verified experimentally then the preference ordering it induces on moral conceptions is uniquely privileged, and therefore in some sense the objectively "correct" one.

Kohlberg himself argues that his data point to six stages in moral development, culminating in reasoning based on "universal ethical principles" that he identifies with what philosophers like John Rawls (1971) have called principles of "justice", i.e. principles by which self-interested agents who viewed each other as having equal rights could come to social agreements that resolve conflicts. This contrasts, for instance, with the utilitarian conception that morality is the "greatest good for the greatest number" (identified in its "rule-oriented" form by Kohlberg as stage 5 thinking), at least on philosophical accounts of this which get quite technical and confusing. Perhaps the greatest difference between the two views is that the latter stage places more emphasis on rights of the individual, and makes little room for violations of personal dignity that would follow if the ends justified the means. This is a variant of a "deontological" theory in philosophical terms, whereas the earlier-stage utilitarian conception is more situational, and tends to see moral choices less as attempts to satisfy constraints of principle and more as attempts to weigh competing values, while recognizing that strict rights themselves have great utility [B]. But it is not the details of these philosophical positions that I wish to discuss in this paper, nor will I be concerned with Kohlberg's claim that an experimentally-verified invariant sequence implies one true morality. Rather I would like to raise several of what I consider to be serious problems with inferring the existence of an invariant sequence for moral development from Kohlberg's data. Even if Kohlberg's philosophical position is correct, in my opinion he is very far from demonstrating scientifically that moral development follows an invariant sequence.

The invariant sequence hypothesis (called by Kohlberg a "postulate" [B], p. 42) holds that some forms of moral reasoning are more "mature" than others in the sense that persons may change during the course of their lives (empirically, until well into adulthood) from one form or "stage" to another, more mature one, but that they do not change in the reverse direction, so for instance a person's reasoning pattern may change from that of a rule-utilitarian (stage 5) to one that applies universal principles of justice (stage 6), but not vice versa. A strong aspect of the empirical claim is that there is an endpoint (stage 6) which is the "cat's meow" of ethics, beyond which there is, in a word, nothing. Kohlberg summarizes the claim of a "formalistic normative theory" (which he advocates) as follows: "Stage 6 is what it means to judge morally. If you want to play the

moral game, if you want to make decisions which anyone could agree upon in resolving social conflicts, stage 6 is it." ([A], p. 218) The mechanism of data scoring employed by Kohlberg reflects this orientation. Kohlberg gives his subjects moral dilemmas or questions, for example, Should a man steal a drug to save his wife's life?, and by an elaborate procedure for scoring, his manual-trained judges classify the subjects' responses into stages according to each stage's "criterion concept", or "the reasoning pattern that is most distinctive of ... [that] ... stage." ([B], p. 45) So every response must be viewed as a manifestation of one of these stages, and in particular there is no provision for moving beyond stage 6. As mentioned above, the claim that there is an endpoint would seem to be one of the things the experiment should test, but Kohlberg assumes it as part of his scoring apparatus, partly on the grounds that (he claims) stage 6 is philosophically ultimate. The lack of a separate empirical test for this claim that there is an endpoint is one possible criticism one can make of his methodology. The remaining four problems that I will discuss in this essay all call into question Kohlberg's claim that the sequence of moral stages is invariant.

1. Sequence is not invariant under aspect-scoring. Although this is perhaps the most difficult criticism to make sharp, it also seems to come the closest to the dominant, though murky, sensation one feels in reading about Kohlberg's procedures that something is not right about them. Kohlberg [B] describes three systems of scoring that had been tried up to 1976, the first of which was called "aspect-scoring" and involved classifying responses in terms of twenty-five "aspects" of the moral stages. In discussing the results, Kohlberg notes that the classifications of data yielded examples of "inversion of sequence", in which, for example, "a small number of individuals 'regressed' from stage 4 to stage 3, or skipped from stage 3 to stage 5." ([B], p. 43) Kohlberg attributes the failure of this scoring system to meet the invariant sequence hypothesis as a failure of aspect-scoring rather than of the hypothesis ("This method turned out, however, to contain too much extraneous content to yield a measure or classification meeting the invariant sequence postulate," p. 42; "These inversions [discussed above], in turn, could be seen as due to an inadequate definition of stage 4, ... As a result, we redefined as stage 3 any law-and-order thinking which did not display a social system perspective," p. 43). Kohlberg subsequently developed two other methods of scoring, the first called "intuitive issue scoring" and the second called "standardized issue scoring". Although the manual for this latter method was still being worked on when [B] was written, Kohlberg indicates that invariance had by then been achieved "in longitudinal data [scored by the 'structural stage method', apparently a synonym for issue scoring]." ([B], p. 46) The philosophy in constructing these systems was the same as that which had led failure of invariance under aspect-scoring to be attributed to the scoring system, namely that the stage definitions and scoring procedures should be modified until the most invariance was observed. This is in fact the criterion of construct validity Kohlberg employs ("our conception of construct validity implies assignment of individuals to stages in such a way that the criterion of sequential movement is met," p. 47). So the game Kohlberg is playing is essentially to construct a set of stage definitions and a scoring system that maximize invariance. This is a very precarious strategy. If the system is sufficiently robust (e.g., if there are no exceptions to invariance) then it will have a great deal of surprise value as a theory because one might not have thought that such a highly predictive psychological theory would be possible. On the other hand, if the theory has exceptions, as we are led to expect from any psychological theory, then he is open to the charge of simply finding that set of stages that is the most invariant, rather than demonstrating invariance. Invariance is a very strict criterion; as soon as there are exceptions, it is disproven.

*This does seem like a peculiar use of construct validity, which usually refers to how well a measure fits not only with other convergent measures, but also with other divergent measures.*

*I'm not sure it has to be that strict. Just a "universality" in linguistics can be seen as a matter of probabilities.*

2. Scorers can cause the data to obey invariance. A second, slightly clearer criticism one might make of Kohlberg's approach is that the scorers have perhaps too many clues about relevant co-variates for the subjects (like their age and intellectual capacity) which, combined with their knowledge of the hypothesis and stages, could result in responses obeying invariance under the scoring procedure. In reading a response that contains many clues that it comes from a very mature, intelligent person, a scorer trying (even unconsciously) to affirm the theory might be much more likely to assign that person to a higher stage. It is very difficult to tell how much the rated stage depends on the content of the response, especially since in ethics a response may be read so many different ways, and how much it reflects the other signs of mature thought (clear expression, vocabulary, logicity) in the text of the response. Thus Kohlberg's claim may reduce to one that says that familiarity with the observed predominant patterns of change in moral reasoning, combined with other clues about maturity, are sufficient to tell how old or mature one is in terms of a scale with six values. If this were true then it would explain the invariance results he seems to have obtained without proving that beliefs about morality progress through an invariant sequence in a way that could be characterized independently of their particular expression. It would be quite a challenge indeed to eliminate the other clues from the data, but it is interesting to wonder whether, if the responses could be recoded in either a neutral language or a language characteristic of a randomly chosen age group for each subject, the scorers could, knowing this, still classify the responses in a way that would obey invariance.

Good point

3. Types of questions asked may unfairly favor one style of moral reasoning. Kohlberg is making the claim about his test procedure that it is a test of "moral reasoning" for some appropriate meaning of this term. But it is very clear that Kohlberg's view of the meaning of the word 'morality' is identical with the principles of stage 6: He says, "Moral judgments, unlike judgments of prudence or aesthetics, tend to be universal, inclusive, consistent, and grounded on objective, impersonal, or ideal grounds," ([A], p. 215) and, "The individual whose judgments are at stage 6 asks 'Is it morally right?' and means by morally right something different from punishment (stage 1), prudence (stage 2), conformity to authority (stages 3 and 4), etc." (p. 216) On this view, conceptions that hold that the right choices as highly situationally dependent and are not reducible to principles are just less moral, by definition. It may well be, in fact, that the ordinary language meaning of 'moral' refers more to considerations of principle than of relative values in the making of social choices, but I am inclined to ask whether a theory of moral or ethical development should reflect this fact about the language. If what we meant by "moral" in so labeling this aspect of development we were trying to get at was the reasoning by which we make choices, then using the more specific meaning of the word 'moral' to justify one form of such reasoning as privileged seems to be to take unfair advantage of the fact that the word closest to (but not exactly) the one we wanted was the word 'moral'. I think that this subtle mistake has deeply penetrated Kohlberg's choice of the test instrument, in that his questions emphasize the principled aspects of choice-making. The questions are worded in a way that favors interpretation in terms of the application of right-wrong principles, by asking in the yes-no form ("Should the husband have done that?", [B] p. 42; "Does it matter whether the wife is important or not?", p. 44) and by asking questions that clearly steer the subject toward an interpretation in terms of principles and rights ("Is it a husband's duty to steal?", p. 42). So one way to interpret his results is that the invariant progression toward justice-based reasoning simply reflects stages in understanding the question. These need not be identical with stages of ethical reasoning more broadly defined. As a thought experiment, we might imagine what types of answers subjects would converge to if they were asked questions that did not effectively ask for deontological

This is an excellent idea (After all if computer programs can assess your writing style for the "grade level" of various sections, it should be possible to translate systems directly from one level to another, and then to assess reliability across levels)

you're right that Kohlberg refers to "moral development" when he means "the development of moral reasoning" (since this is all that he is actually assessing). He acknowledges that moral behavior might not be included with moral reasoning; however, by reducing "moral dev. to reasoning" he implies that behavior is not relevant to moral development. This does not seem like a very satisfying implication

answers, but instead asked them to weigh consequences, as a utilitarian would. As an example question, we might ask, "If a few people had to be killed in order to prevent a nuclear war, would the latter consideration outweigh the violation of rights entailed by the killing?" In longitudinal data we might see subjects converging on the utilitarian, rather than the deontological, answer, just because they are being asked to weigh the consequences rather than to ask whether something "should" be done, or whether it would violate anyone's rights. The sensitivity to question wording and to the types of dilemmas posed could be quite profound, and does not seem to be addressed by Kohlberg.

Why does  
it follow?

4. Sequence of development may be culturally determined. Kohlberg sees the movement toward stage 6 as an inevitable consequence of the development of logical (e.g., formal operational) thought. But unless his philosophical argument can be made more convincing, we are faced with the need to verify this inevitability in some other way. As it stands from the presentation in [B], Kohlberg's invariance, even if we grant that his questions capture the morality we are trying to test for (see above objections), could be explained as an artifact of Western culture. If Americans really do move toward a deontological ethics based on justice, this might be just the result of socialization, living as we do in a land where it is written that we are endowed by our creator with "certain inalienable rights", and where there is a Bill of Rights. An essential test of the invariant sequence hypothesis would be whether it holds in a country like the Soviet Union, where the concept of rights is certainly de-emphasized. Might not there have been, in Nazi Germany, some Germans who changed from believing that people should not be killed for the "good of the nation" to believing that they should be. Examples of this kind would pose severe problems for Kohlberg.

And Japan, where "local" bylaws are of enormous importance in guiding action, and where "logic" "reason" "principles" etc are regarded with deep suspicion at the highest levels of thought, e.g. in Buddhist thought.

To briefly summarize, then, I think that it is very unlikely that invariance in the progression of reasoning about social choices can actually be demonstrated, primarily because reflecting on it makes me doubt that invariance exists in that sense. The field of moral development is an exceedingly difficult one to think about in a precise way, and the philosophical arguments are almost unbearably abstruse. I think, though, that before the philosophical implications of Kohlberg's approach can be fully comprehended, the problems involved in attempting to demonstrate invariance experimentally need to be more carefully worked through.

#### REFERENCES

Kohlberg, L. (1971) "From Is to Ought: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development". In Mischel, T. (ed.) Cognitive Development and Epistemology. New York: Academic Press, pp. 151-235.

Kohlberg, L. (1976) "Moral Stages and Moralization: The Cognitive-Developmental Approach". In Lickona, T. (ed.) Moral Development and Behavior: Theory, Research and Social Issues. San Francisco: Holt, Rinehart and Winston, pp. 31-53.

Rawls, J. (1971) A Theory of Justice. Cambridge, MA: Harvard University Press.

A  
A very thoughtful  
and interesting paper.