The Moral Judgment Test: Comments on Villegas’ Critique

Georg Lind

University of Konstanz

Contact information:

Prof. Dr. Georg Lind, Department of Psychology, University of Konstanz, 78457 Konstanz, Germany, or e-mail (georg.lind@uni-konstanz.de).
Summary. — The Moral Judgment Test has been designed for research and evaluation studies (but not for individual diagnosis). It is an objective, reliable and very valid test of participants’ ability to make moral judge arguments. It has been designed on the basis of a multivariate experimental and information processing paradigm rather than a conventional psychometric paradigm. Hence “reliability” and “validity” needed to be operationalized in a new, and also more rigorous way. Neglect of its intentions and of the supporting data renders its critique as inadequate.

In the past, moral and democracy education programs have mostly been evaluated through tests of moral and democratic attitudes or tests of declarative knowledge, that is, of moral and democratic concepts. However, we were not able to study the impact of education on moral and democratic competencies, and to evaluate methods of teaching them. The Moral Judgment Test (MJT) by Lind (in press; Lind & Wakenhut, 1985) has been designed to fill this void. Because the MJT is highly reliable, valid and economical, it has helped to evaluate and improve methods of moral education such as dilemma discussion (Lerkiatbundit et al., 2004; Lind & Althof, 1992; Lind, 2002; 2003) and teacher education.

Villegas’ (2005) critique of the MJT is based on a false understanding of the MJT and is a rather biased report of the cited literature. Readers who want to make a reasoned judgment on the MJT are referred to the original literature (Lind, in press) and the author’s website: http://www.uni-konstanz.de/ag-moral/ This article is to summarize the 30 years of research.

The trait which the MJT is to measure is moral judgment competence. This is defined by Kohlberg (1964), as the “the capacity to make decisions and judgments which are moral
(i.e., based on internal principles) and to act in accordance with such judgments” (p. 425).

With the MJT, the participant is tested whether he applies his or her moral ideas to the judgment of other people’s arguments rather than judging them merely on the basis of their agreement with his or her own opinion. The MJT also follows Kohlberg’s requirements for a new structural measurement method: “I define stages solely in terms of cognitive structures, or ways of thinking or judging” (Kohlberg, 1984, p. 398). “The structures we seek to tap in test construction and arrive at in test scoring are abstractable from responses as their form or quality” (p. 402). “[Structure] is a construct rather than an inference, and is warranted only on the grounds of ‘intelligible’ ordering of the manifest items. One might say that the hypothetical structure is the principle of organization of the responses” (p. 408). These requirements are at odds with classical psychometric paradigm of attitude measurement and testing, because this paradigm interprets response consistency as a property of the test, i.e., “reliability” and “validity,” and thus does not allow the assessment of “structures,” “forms of response,” or “intelligible ordering of the manifest items” (Lind, 1989).

Like Kohlberg (1984), we interpret response consistency as a rich source of information on a participants’ cognitive functioning and, therefore, use methodological paradigms like multivariate experimental N=1 design (Kelly, 1965), information integration theory and cognitive algebra (Anderson, 1990), and facet analyses and structural theory (Guttman, 1971). Hence, we have operationalized “reliability” and “validity” accordingly (Lind, 1985; in press; Lind & Wakenhut, 1985).

Because the MJT is objective (closed questions and scores derived by computer algorithms), and because it is fully standardized, it is fully reliable. No subjective rating is applied that can bias the test scores toward some end, nor are the items selected to artificially maximize stability and variability of scores. Of course, because the MJT was to be used as an
instrument for measuring developmental and educationally induced change, it must not produce stable scores, nor must it measure changes due to simulation. In fact, the MJT shows high short term stability of $r = 0.90$ (Lerkiatbundit et al., 2004), and resistance to upward simulation (Lind, 2002). Yet it also shows high sensitivity to educationally induced improvement of moral competencies (Lind, 2002; in press).

For validating the MJT, Lind (in press) set up five unprecedentedly rigorous empirical criteria, which are grounded in cognitive-developmental theory:

1. Rest’s (1969) “preference hierarchy,” which postulates that the preference for each of Kohlberg’s stages of moral orientation, should be ordered exactly in the way Kohlberg had predicted: Stage 1 is preferred least and the higher stages are preferred increasingly according to their number.

2. Kohlberg’s (1958) “quasi-simplex” of stage-preference inter-correlations, which means that the preference for neighboring stages should highest and for more distant stages should be lowest so that the correlations are highest along the diagonal of an inter-correlation matrix and fall off toward the corners.

3. Piaget’s (1976) “affective-cognitive parallelism,” that is the prediction that the preferences for each Kohlbergian stage should be systematically correlated with the index for moral judgment competence in such a way that the higher the judgment competence the more should lower stages be rejected and upper stages accepted.

4. Lind’s (2002) “education theory of moral development,” which means that the amount of good quality education experienced by the participant, should be strongly and positively correlated with the index of moral judgment competence.

5. Finally, Emler et al.’s (1983) test of upward-simulation should be negative, that is, participants should not be able to simulate the index of moral judgment competence upward
through simple instruction, as they can when instructed to simulate tests of moral attitudes (Lind, 2002).

In regard to all five criteria, the MJT has shown to be exceptionally valid in dozens of validation studies conducted worldwide, in spite of the fact that the items of the MJT have not been submitted to item-analysis nor item-selection. After 30 years of research on, and use of, the Moral Judgment Test in many studies worldwide it can be clearly recommended for research and evaluation studies.

References


