

The Dual Aspect Model of Moral Behavior –
An Experimental Test of Piaget’s Theory of Affective-Cognitive Parallelism¹

Georg Lind
University of Konstanz²

Introduction

Half a century ago, Jean Piaget suggested a new approach to the study of moral (and non-moral) behavior which we have called *Dual Aspect Model*, and which, as we shall see, has opened up a new, very prosperous field of research into the nature and development of human morality. When I came across Piaget’s suggestion, in the early 1970s, I was thrilled. I imagined that if this notion, which has received little attention in research on moral behavior and development, were fully understood, it could revolutionize psychology and education. I felt that this theory could solve some of the most troubling problems of moral psychology, foremost the definition and measurement of *moral competence*, a notion which Lawrence Kohlberg (1964) had just introduced. The central postulates of Piaget’s new approach are:

“The two aspects, affective and cognitive, are at the same time inseparable and irreducible” (Piaget & Inhelder, 1969, p. 158; emphasis added).

“Although cognitive and affective factors are indissociable in an individuals’ concrete behavior, they appear to be different in nature” (Piaget, 1981, p. 3).

“Sentiments express the interests and values of actions, intelligence constitutes the structure” (Piaget, 1951, p. 220).

To my knowledge, Piaget did not develop a specific testable hypothesis or measurement method on the basis of his dual aspect model. Yet in several of his writings he suggested a general hypothesis, namely he predicted that both aspects would be closely correlated: “We [...] find a marked parallelism in their respective evolutions” (Piaget & Inhelder, 1969, p. 21). “*We shall be able to put intellectual structures and the levels of affective development in*

¹ Revised version of my invited address to the 2nd International Colloquium on Genetic Epistemology and Psychology: Current Interlocutions and Discussions held from Nov. 7 - 10, 2011 in Marilia, Sao Paulo State, Brasil. To be published in the proceedings of the Colloquium.

² Contact: Prof. Dr. Georg Lind, University of Konstanz. E-mail: georg.lind@uni-konstanz.de. Website: <http://www.uni-konstanz.de/ag-moral/>

parallel, stage by stage” (Piaget, 1981, p. 12).

In the past three decades my colleagues and I have used the Dual-Aspect-Theory to develop a new approach to the measurement of affective and cognitive aspects of moral behavior as distinct yet inseparable aspects. We also explicated Piaget’s assumption of an affective-cognitive parallelism. We hypothesized that the higher people’s moral competence (cognitive aspect) is, the more pronounced they prefer higher order moral orientations, and the more they reject lower order moral orientation. We have shown elsewhere that Piaget’s notion has powerful implications for moral psychology and education (Lind, 2002; 2009; 2008 a; 2008 b; 2010 b). In this paper I want to show that the Piagetian prediction can be experimentally corroborated.

The slow progress of “soft psychology”

In order to understand the revolutionary implications of Piaget’s notion of affect and cognition being two inseparable but distinguishable aspects for moral psychology (and beyond), we need to review the hidden assumptions underlying mainstream psychology. These hidden assumptions, it seems, are so powerful that they have prevented us for a long time from understanding the difference between aspects and *components* of behavior and the importance of this distinction for the progress of psychological research and educational practice. Maybe the importance of this distinction would be more readily accepted if the reasons for the slow progress of “soft psychology” (Meehl, 1978) were better understood.

In spite of a vast amount of experimental studies, the science of psychology is making only slow, if any progress. As Wittgenstein (1953)³ argued, this is not so much because of a lack of experimental and statistical methods but because of a lack of conceptual parsimony and clarity. In psychology ‘complexity’ is adored, not simplicity. Most psychologists seem to assume that all behavior is complex. However, complexity lies in the eyes of the beholder. Anything that we do not yet understand appears to be complex. But once we understand things they can look quite simple. It seems that our deep belief in complexity makes most of us rest content with *not* understanding psychological processes.

³ “The confusion and barrenness of psychology is not to be explained by calling it a ‘young science’; its state is not compatible with that of physics, for instance, in its beginnings. ... For in psychology there are experimental methods and conceptual confusion. The existence of experimental methods makes us think we have the means of solving the problems which trouble us; though problem and methods pass one another by.”

In the Dark Ages of medieval times, natural science was handicapped by a similar thinking. This handicap was overcome only when it was discovered that theories about nature must not be confused with nature itself, and the highest virtue of a good theory was parsimony of concepts. One of the pioneers of this new thinking was William of Occam (1285 - 1349). “Occam’s razor” became proverbial: Entities should not be multiplied without necessity (“*entia non sunt multiplicanda sine necessitate*”). Rigorous striving for simplicity has made modern science and technology possible. Only if predictions are simple (not simplistic!) and clear can they be tested for information value and empirical truth (Popper, 1968). Only knowledge which can be presented in the form of a simple theory can be taught and can be transformed into powerful technologies.

Moral psychology is still far from this ideal. Consider, for example, this influential definition of the new concept of “moral foundations:” They “are not values or virtues. They are the psychological systems that give children feelings and intuitions that make local stories, practices, and moral arguments more or less appealing during the editing process. [...] The foundations are the main ‘evolved psychological mechanisms’ that are part of the ‘first draft’ of the moral mind” (Graham et al., 2009, p. 1031). It is difficult, if possible at all, to imagine how this complex concept can be observed or measured in an objective way. The authors saw no other way to resolve this problem than through measurement by fiat: “The items were written to be face-valid measures of concerns related to the five foundations” (p. 1032).⁴

(Wittgenstein, 1953, chapter xiv)

⁴ The problems which we discuss here are not confined to moral psychology but trouble psychology and the social sciences as a whole. As Jane Loevinger (1976) noted, “many methods [of measurement] are flawed. They do not coordinate concept type with scoring algorithm, they confuse distributions with profiles, and they use rules and parameters that are arbitrary to the point of caprice. For measurement to serve as the leading edge of a scientific discipline, it must be informed by theory, and there must be clear lines for results to feed back as corrections to theory. Most of the approaches have inadequate articulation of theory, method, and data” (p. 240). Similarly the social psychologist W. A. Scott (1968) argues that “perhaps the most influential, and certainly the best developed source is psychometric theory, or the theory of mental tests (for example, Gulliksen, 1950). Though currently under fire for its inadequacies [...], it at least has the virtue of explicitness, which renders its inadequacies obvious” (p. 208). George A. Miller (1969), former president of the American Psychological Association noted: “What is lacking is a psychological theory that dictates explicitly which items should be included on the test. Then the criterion would be used, not to validate the test, but to validate the theory on which the test was based. Such an explicit theory – if it were true – would resolve all doubts as to whether or not the test actually measured what it was intended to measure. Questions of validity would be transferred to the larger domain of psychological theory in general, and the tests would become an instrument of

Yet if there is no clear and consistent relation between concept and method of observation or measurement, there is no valid base for testing theories and for reaching an agreement on the truth of theories on the basis of scientific research. Truth remains a matter of social power and status. We can test the empirical validity of concepts only if the method of observation is consistent with these concepts, that is, if our measurement is *theoretically valid*. If the validity of our measurement is unknown, we cannot be sure whether the data we get from measurement falsify our theory or the measurement. Not only the development of good theories depends on good measurement, but good measurement depends also on good theories. Good theories means that we clarify and simplify our theoretical concepts, that is, that we clarify what our concepts mean empirically and rid our concepts of undefined terminology and unnecessary assumptions. Otherwise, we cannot know whether our theory or our method is wrong.

In moral psychology, Piaget has contributed much to the clarification of the two concepts of moral affect and moral cognition in his writings, especially in *Moral Judgment of the Child* (Piaget, 1965) and his article on the measurement of the cognitive-structural aspect (Piaget, 1971): “Structures [...] are expressed in regular forms of responses that we believe we are discovering in the subject's behavior. We also feel that if the underlying structures did not exist, we would not be able to explain such behavior. But the subject is not aware of these structures. He is not a professor of psychology. ... He simply uses them” (p. 3).

Similarly Kohlberg, in his Heinz Werner Memorial Lectures, discussed at length the problem of coupling concept and method of moral psychology (Kohlberg, 1984): “One who seeks to locate responses with regard to underlying structure makes a distinction between 'achievement' and 'process' [...] In order to arrive at the underlying structure of a response, one must construct a test, [...] so that the questions and the response to them allow for an unambiguous inference to be drawn as to the underlying structure. [...] The test constructor must postulate structure from the start, as opposed to inductively finding structure in content after the test is made. [...] If a test is to yield stage structure, a concept of that structure must be built into the initial act of observation, test construction, and scoring; it will not emerge through pure factor-analytic responses classified by content” (pp. 401 - 402). Already at the outset of his research into moral behavior and development, Kohlberg (1958) postulated that

research comparable in power and dignity to experiments conducted in the laboratory. [...] The current trend [...] is toward the explicit use of psychological theory in constructing new tests. With this changed emphasis the psychometric problem enters a new phase, less technical and more scientific” (p. 369).

“a moral act or attitude cannot be defined either by purely cognitive or by purely motivational criteria” (p. 16; emphasis added, GL).

Yet, in spite of these prominent voices, moral psychology is still plagued with the lack of agreement between method and concept. Most psychologists still seem to believe that the arduous process of defining and clarifying concepts like affect and cognition in the moral domain can be short-cut through statistical methods. In their seminal study into the nature of deceit, Hartshorne and May (1928) argue: “If [the test] can be shown to be reliable, then it is ipso facto a valid measure of the particular behavior in question in the particular types of situation embodied in the test.” In a footnote they define validity as “the square root of its reliability” (p. 142). Similarly, Kohlberg (1984) argues that “from the point of view of stage theories like Piaget’s or mine, test reliability and test construct validity are one and the same thing” (p. 424). This is of course false. Obviously these authors confuse the statistical concept of *reliability*⁵ or *consistency* with psychological consistency. Statistical consistency of responses of a *sample* of participants is something totally different from the consistency of an *individual*’s answers. In a sample the rank order of individuals regarding their test scores can persist over some time (which would show in a high reliability coefficient) and still each individual may respond rather inconsistently. Neither must the consistency of test-scores be confused with the theoretical validity of a test of moral attitude or competence. We can measure an object very reliably and not know what we are measuring.

Some moral psychologists argue that modern statistical models like Item-Response-Theory define and clarify the objects of measurement. However, these models also rest on questionable psychological assumptions that are hidden in allegedly methodological decisions. The IRT rests, among other things, on the assumption that all responses to test-items are distributed in the same way (usually “normality” is assumed), and that responses are linearly related with the “underlying trait”. If the data do not fit this expectation, many test-constructors exchange the test-items which do not seem to “work.” In other words, “the measurer has cheated the validity test by using empirical evidence to modify the [...] construct”

⁵ Reliability can mean that participants’ scores on a test taken at one point of time correlate very closely with their scores at another time (“re-test reliability”); in other words that, for example, they do not develop morally, or that they all develop at the same rate so that their ranking on the test remains perfectly stable. It is hard to see how this is related with the question whether a test really measures what it purports to measure. A different meaning of reliability is that all participants respond to all items of a moral test in a consistent way (whereby ‘consistent’ may be defined in various ways).

(Wilson, 2005, p. 161). Originally, the assumption of normal (bell-shaped) distribution was developed for the statistical treatment of errors of measurement. Yet, psychologists adapted this concept prematurely to the distribution of human traits. In her very influential textbook on personality, Anastasi (1958) postulated: “Many of the distributions found in differential psychology likewise fit the mathematical specification of a normal curve, especially when they are obtained through the use of carefully constructed measuring instruments with large representative groups” (p. 28). Similarly, the moral psychologists May and Hartshorne (1926) rely on this assumption: “All we can do is to fall back on the normal curve and use the SD [standard deviation] as our unit” (p. 153), though they bear some caveats in mind (see below). Sprinthall et al. (1994) also argue that “many behavioral measures in educational psychology conform to what statisticians call the normal curve. [...] So many measurements come so close to this ideal that it is of utmost importance” (pp. 435 - 436). None of these authors provide experimental evidence for their belief. For many psychologists normal distribution of human traits has become a religious belief.⁶

Yet this belief is wrong. Only the distribution of measurement *error* is ‘normal or bell-shaped,’⁷ not the distribution of moral or other traits. In their seminal experimental studies into the *nature of deceit*, May and Hartshorne (1926) concede: “It may be questioned [...] whether we are justified in assuming the normal distribution of such a tendency as dishonesty and basing a scale on this assumption” (p. 147). Moreover, “an investigation of the distributional characteristics of 440 large-sample achievement and psychometric measures found all to be significantly non-normal at the alpha .01 significance level” (Micceri, 1989, p. 156). If our measurement results in normally distributed data we should not see this as confirming the validity of our data but as a hint that our measurement has produced mostly error.

Also the belief of test-psychologists that all (latent) traits would be mono-causally related to the measurement data is wrong. Observed behavior is hardly ever, or never, solely the function of the mental states of the observed person, nor is it solely a function of one isolated mental state. Most, if not all, observed behavior is also a response to situational variables, foremost a response to the observation itself. In interviews and tests, the participants’ reaction

⁶ “It is a fortunate coincidence that the measurements of many variables in all disciplines have distributions that are good approximations of the normal distribution. Stated differently, ‘God loves the normal curve!’ ” (Hopkins & Glass, 1978, p. 95)

⁷ “I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the ‘Law of Frequency of Error’ ” (Galton, 1889, p. 66)

is not only determined by the targeted trait but also by the wording of the questions, by certain traits of the interviewer, by the perceived purpose of the interview, by the participants' desire to be socially accepted, and by many other features of the person and the interview-situation. Wuttke (2007) has shown that test-data in school achievement tests are determined by more than one factor. The same is true for tests of moral judgment behavior (Lind, 1978; Anderson, 1991). The co-determinants of responses are not random ("measurement error") but are mostly systematic and thus part of participants' personality and of the experimental situation. Therefore, the alleged characteristics of tests like 'reliability' and 'validity' are as much a characteristic of the participants who supplied the data than of the test itself. The conventional attribution of response inconsistency to the test is a matter of belief rather than of scientific reasoning (Lind, 2010 c). The false attribution of consistency to the test and the test items is immunized against falsification by statistical methods of "item analysis" and "item selection." If test-scores do not seem to be consistent enough, one tries to identify the items which are to be blamed for this, and replace them with new items which produce a better fit of the response pattern with the expectation that all behavior is consistent, thus immunizing dubious theories against falsification.

Neither Piaget, nor Kohlberg, nor many neo-Piagetian and neo-Kohlbergian researchers are free from such 'saving circularities' either. When Piaget (1965) set out to study *the moral judgment of the child* he believed that ideally this should be done through direct experimental observation (p. 115). Yet he argues that "you cannot make a child act in a laboratory in order to dissect his moral conduct" (p. 112). Hence, "it is the moral judgment that we propose to investigate, not moral behavior or sentiment" (p. 7). He argues that this methodological shift would not impede the validity of his data; he assumed that "the things that children say to us constitute, as compared to their real conduct, a conscious realization or a 'derivation' reflection" (p. 115). To support this belief, Piaget points at "the results of our method [which] are relatively constant and, above all, they evolve with a certain regularity according to age" (p. 120). However he concedes that his belief can be challenged: "But there may also be no connection whatever between the two. On this view, the child's moral theories would be mere chatter, unrelated to his concrete evaluations. [...] It may be for the benefit of the adult rather than for his own use that the child gives his answers" (p. 117). Moreover, Piaget was aware of the multiple determination of test responses: "Verbal evaluations made by our children are not of actions of which they have been authors or witnesses, but of stories which have been told to them. The child's evaluations will, therefore, be verbal, as it were, to the second

degree” (p. 119). Indeed, correlation with age would be a sign of validity of measures of morality only if we could be sure that there is no stagnation and no regression. But we can test this assumption only if the measurement has not been immunized against regression through item selection and scoring.

Even modern ‘cognitive neuroscience’ suffers from the same lack of coupling concept and method and from a Pre-Occamite multitude of undefined entities. As the neuroscientist Sharma Borg (2008) observes: “While this classic distinction has been useful in initial efforts to sort out the neural underpinnings of sociomoral behavior, the terms ‘reason,’ ‘emotion,’ ‘intellect,’ or ‘cognition’ are commonly used in the scientific literature without defining what they are supposed to mean [...] Perhaps the emerging field of moral neuroscience hasn’t defined the terms ‘emotion,’ ‘reason,’ ‘cognition,’ ‘intellect,’ etc. because we aren’t yet sure what they are” (pp. 161-162). Perhaps because of this conceptual confusion, some eminent cognitive (!) neuroscientists believe that moral behavior can be explained without reference to cognition (Haidt, 2001; Greene, 2008; Graham et al., 2009). Unfortunately, neuroscience has developed largely disconnected from the large body of Piagetian and Kohlbergian moral psychology. The exceptional study by the neuroscientists Kristin Prehn and her colleagues (2008) lets us sense how fruitful a cooperation between both paradigms of moral science could be.

Component Approaches in Moral Psychology

Moral psychology, like main stream psychology, has embraced, and still embraces, component models of human behavior. Component models are deeply entrenched in our research practice as well as in our educational institutions. They entail certain ways of moral measurement (separate instruments for each component) and certain moral classroom practices (separate methods for the various components). Some describe the components explicitly. They single out cognitive, affective, behavioral and other components, describe ways of measuring these components separately, and propose educational methods of fostering them separately. Others do not state explicitly their theoretical model or even name its elements’ “aspect” even though the methods of measurement and education that they suggest are obviously based on a component model, because they describe them as *separable*.

As we have already seen, even Piaget often leans toward a component model. He talks

about affect and cognition as if they were separable. “Affective life, similar to intellectual life, is continuous adaptation, and both of these adaptations are not only parallel but interdependent [...]” (Piaget, 1951, p. 220). “We shall be able to put intellectual structures and the levels of affective development in parallel, stage by stage” (Piaget, 1981, p. 12). “Parallelism between intelligence and affectivity would require that analogues of conservation and operations be found in the affective domain” (p. 13). No doubt, Piaget was convinced that affect and cognition, morality and intelligence were closely related, but his treatment of them implies that he thought of them as separable components. In the table of developmental stages, he listed them side by side as separable kinds of behavior (Piaget, 1981, p. 14).

In contrast, for Kohlberg affect and cognition were so closely related that he even did not distinguish them clearly in his stage theory of moral development. “A systematic general observation of moral behavior, attitudes, or concepts in terms of such a set of formal criteria of morality [...] cross-cuts the usual neat distinctions between moral knowledge or beliefs on the one hand and moral behavior or motivation on the other, since a moral act or attitude cannot be defined either by purely cognitive or by purely motivational criteria” (Kohlberg 1958, p. 16). However, as we have also seen, Piaget argued that inseparable does not mean indistinguishable. For example, we cannot separate the size of a ball from its weight because both are aspects or properties of the ball, not separable components like air and rubber. Yet we can clearly distinguish the aspects “size” and “weight” and can measure them independently. Analogously, we should be able to distinguish affective and cognitive aspects and measure them independently, even though we cannot separate them nor can we measure them separately (Lind, 1978; 2008 a; 2010 b).

Component models of moral behavior are en vogue in current moral psychology. Rest (1984), who considers his theory “Neo-Kohlbergian,” suggested a four component model (cf. also Rest et al., 1999). Apparently he supposes that these components can be separated from each other. Consequently separate tests have been proposed for measuring each component. From the point of view of neo-Piagetian and neo-Kohlbergian theorizing, this is a questionable assumption. As Higgins (1995) points out: “However, one should note that there are cognitive aspects to all of Rest’s components, and Kohlberg’s idea of a stage as a structured whole or a world view cuts across Rest’s component model” (p. 53). Rest’s component model is clearly at odds with Piaget’s aspect-model.

Similarly, neuroscientists often use a component-language when talking about the relationship between emotion/affect and cognition (cf. Haidt, 2001; Greene, 2008). Yet, as

Greene (2008) observes, “often ‘cognition’ is used in a [...] sense that contrasts with ‘emotion,’ despite the fact that emotions involve information processing” (p. 40).⁸

Moral Competence

Moral competence is broadly defined as the ability we need to apply our moral ideals in every-day life, especially to resolve moral conflicts. This definition is derived from Kohlberg’s definition of *moral judgment competence* as “the capacity to make decisions and judgments which are moral (i.e., based on internal principles) and to act in accordance with such judgments” (Kohlberg, 1964, p. 425). Initially, Kohlberg wrote that moral (judgment) competence could be studied through observing children when they were confronted with a (difficult) moral task: “We felt that it would be easier to analyze qualitatively a case in which the situation demanded more than a child could respond to than to analyze a case in which a child wanted more challenge than the situation could provide” (Kohlberg, 1958, p. 76). For this purpose he developed a new method of assessment, the *Moral Judgment Interview* (MJI). Like Piaget, he confronted children with stories. Initially he was not interested in children’s theories about right or wrong but in the structure of their behavior. He probed into participants’ moral judgment competence by asking why-questions and by confronting them with counter-arguments. “The responses of subjects to the dilemmas and their subsequent responses to clinical probing are taken to reflect, exhibit, or manifest the structure” (Kohlberg, 1984, p. 407).

Later Kohlberg and his colleagues lost sight of this structural approach to the measurement of moral judgment competence, when they changed their measurement method in order to save a core assumption of cognitive-developmental theory from refutation (Lind, 1989). When Kohlberg and Kramer’s (1969) longitudinal study produced cases of regression, Kohlberg and his associate did not accept these as a refutation of their postulate of “invariant sequence” but argued that the regressions were due to measurement error. Therefore, he and his associates created a new methodological postulate, namely that “the validity criterion of moral judgment development is [...] that of an organization passing through invariant stages” (Kohlberg, 1984, p. 194). “Before you try to explain data of change and development with a

⁸ In fact, it is hard to envision an “emotional dog which wags with his cognitive tail” (Haidt, 2001) without cognition that controls its tail wagging.

cognitive-developmental theory, make sure your data can be observed with a measure you have made up to fit the sequence rule” (p. 424). After modifying the scoring of the MJI, “as one would expect of a developmental variable, our data show a clear relationship between age and moral judgment stage. The correlation between age and MMS was .78” (Colby & Kohlberg, 1987, p. 47).

However, this success was achieved at the expense of the scientific dignity of cognitive-developmental theory. The postulate of invariant sequence has become immunized against refutation. As Popper (1968) argues, a theory which cannot be refuted by data is pre-scientific at best. And the success did not last long. Studies clearly show that moral competence can regress (Helkama et al., 2003; Lind, 2000; Lind, 2002). Although many Kohlbergians still argue that tests of moral developmental must correlate highly with age in order to be accepted as a valid measure, some chief proponents have silently given up this criterion of test validity (e.g., Colby 2008, p. 393).

But this is not the whole story. As I have shown elsewhere (Lind, 2010 a), regressions can be better explained if we adopt the dual aspect model, that is if we measures the two aspects simultaneously but not in a confused way. In Kohlberg’s interview method both aspects are befuddled, even though in his methodological writings he distinguished the two aspects. On the one hand, Kohlberg defined “stages solely in terms of *cognitive structures*, or ways of thinking or judging” (Kohlberg, 1984, p. 398). “The primary theoretical definition of *structural* moral development is that of an *organization* passing through invariant sequential stages” (Kohlberg, 1976, p. 44; emphasis added, GL). On the other hand, Kohlberg also defined his “stages” in terms of moral affects, i.e., moral orientations, attitudes and motivation. “The present research deals with the interrelated development of basic moral concepts and attitudes” (Kohlberg, 1958, p. 1; first sentence!) He states that the six (or five) stages of moral development could be characterized by different moral orientations underlying thinking and behavior (Kohlberg, 1976). At another place he argues “that the judgment of whether an act is morally right or good, morally bad or wrong, or morally neutral can be decided only by studying the moral judgments and motivations which inform it” (Kohlberg, 1984, p. 393). Yet, in spite of the fact that he saw moral behavior and development as defined by two different aspects, he proposed only *one* index for moral development (“stages” or “MMS”), confounding the two aspects. “By this definition [...], subjects at each higher stage were more likely to act morally in that they were more likely to make judgments of responsibility consistent with their deontic choice and to act on this judgment” (Kohlberg,

1984, p. 523). For example, a participant's responses are scored as "Stage 6" only if she shows strongest preference for Stage 6-type moral reasoning *and*, at the same time, does so consistently across situations.⁹ Because of this confounding, the stage score goes down if the scores on either aspect goes down. The confounded measurement does not let us say which one. If we measure both aspects un-confoundedly (not separately!), as the MJT allows us to do, we find regression only in regard to moral competence (cognitive aspect), not in regard to the affective aspect. Moral orientations remain largely stable, they do not regress (Lind, 2010 a).

Piaget's Affective-Cognitive Parallelism Hypothesis

Piaget (1981) hypothesized that affective and cognitive aspects of human behavior develop parallel.¹⁰ "We shall be able to put intellectual structures and the levels of affective development in parallel, stage by stage" (p. 12). "Affective life, similar to intellectual life, is continuous adaptation, and both of these adaptations are not only parallel but interdependent, since sentiments express the interests and values of actions, intelligence constitutes the structure" (Piaget, 1951, p. 220)¹¹

Piaget not only thought that the two aspects are functionally related but also that they are developmentally parallel. "We shall be able to put intellectual structures and the levels of

⁹ This requirement might explain why in MJT-studies only very few participants are found with Stage-6 scores. Some of the dilemma-stories in the MJT do not seem to require Stage-6 type moral orientations, restraining the scores to the first five "stage"-types not because participants cannot reason on higher stages but because of the method of measurement used (cf. Lind, 1989).

¹⁰ "[Piaget] spoke about affectivity in a broad sense as the energetic source on which the functioning of intelligence depends, drawing the analogy of affectivity as the fuel that makes the motor of intelligence go. [...] need, interest, effort, [...] attraction" (DeVries, 1997, p. 6).

¹¹ Piaget seems not to be immune against confusing the aspect model with the component model, at least implicitly. Talking of "affective life" as distinct from "intellectual life" could be easily mistaken as implying separate components of life. Or read this: "We have assumed that affective decentering is a correlative of cognitive decentering, not because one dominates the other, but because both occur as a result of a single integrated process" (Piaget & Inhelder, 1969, p. 26) Using the term "affective decentering" implies that affects are substances that can spread out, and using the term "cognitive decentering" could mean that decentering and cognition are something different. They are not. Decentering, as Piaget says elsewhere, *is* a way of describing cognitive properties of the human mind.

affective development in parallel, stage by stage” (Piaget, 1981). What developmental parallelism would mean concretely is much debated in research literature. Kuhn et al. (1977) argue that both aspects of moral development are parallel but that cognitive (logical) development always precedes moral development. On the basis of her studies, Nunner-Winkler (1989) discards Piaget’s notion of parallelism, while Lind (2002) defends Piaget’s assumption, with the modification, however, that, in contrast to Kuhn et al.’s theory, we should assume development to be circular: First, the child needs to develop a strong, yet cognitively undifferentiated moral sense that guides and energizes the development of his moral-cognitive structures, which then in turn helps the child to develop more differentiated and integrated moral orientations (Lind, 1989; 2002; 2010 b). Yet, in this paper we will have to limit ourselves to the hypothesis of correlations parallelism.

Corollaries of Piaget’s parallelism hypothesis can be found in the works of Kohlberg and Rest. Kohlberg’s (1958) wrote at the beginning of his dissertation that his “research deals with the interrelated development of basic moral concepts and attitudes” (p. 1). For him “it seemed to be a fact that quantitative consistency in the type [of moral orientation] was associated with qualitative extremeness in expressing its underlying 'principle' ” (p. 94). Even more elaborated is Rest’s formulation of the parallelism hypothesis: “[T]here is evidence that higher moral judgment scores reflect greater capacity and are not merely differences in preference. Studies of moral comprehension indicate that those subjects with higher moral judgment scores also have higher moral comprehension scores, and that subjects with lower moral judgment scores have lower comprehension” (Rest, 1988, p. 188). In this statement, the cognitive aspect is defined as *comprehension* of moral arguments made by others, and the affective aspect is defined as the *preference* for post-conventional moral reasoning. Similarly, Montada (1993) argues that there is a functional link between both aspects by hypothesizing that moral emotions presuppose perceptions or ‘cognitions’ of situations. “These cognitions do not need to be reflected or objectively true, nor do they need to be verbalized or conscious. Nonetheless, they are functional for the arousal of [moral] emotions.” (p. 272).

Today, half a century after Piaget, contemporary moral neuroscience has re-discovered affective-cognitive parallelism, obviously without being aware of Piagetian moral psychology: “The ventromedial and dorsopateral systems do not typically act in isolation, however, and neither do whatever we instinctively mean by 'reason' and 'emotion;' they act in parallel and with constant interaction” (Oliveira-Souza et al., 2008, p. 162).

Piaget quoted much supportive evidence, yet he did not explicitly state how both aspects

and their elements should be measured, nor did he design or conduct experimental studies in order to test this general hypothesis. In particular he did not solve the self-imposed problem of operationalizing the two aspects as aspects and not as components, that is, of designing an instrument which lets us measure both aspects independently as aspects of one and the same pattern of behavior, but not with separate tests of separate behaviors as has been done, for example, in the studies by Kuhn et al. (1977) and Rest (1984).

Therefore, in order to put the hypothesis of correlational parallelism to the test, we need to state it more precisely. We predict that, if Piaget's hypothesis is true, moral affect and moral cognition relate in such a way that *people will prefer more adequate moral orientations, and reject inadequate moral orientations, the higher their ability to make moral judgments*. We should, however, keep in mind that this prediction will only be confirmed by empirical studies if the participants have no reason to fake their moral orientations upward (e.g., if they have good reasons to believe that the test is for high stakes), or if the participants lack motivation to perform at their actual level of moral competence. In both cases, the prediction may still be supported by the data but not as clearly as we should expect.

Method

To test Piaget's hypothesis of affective-cognitive parallelism (that is, our more specific reformulation of this hypothesis), we used the *Moral Judgment Test* (MJT) (Lind, 1978; 2008 a). In line with Piaget's concept of distinct-but-not-separable aspects, we have designed the MJT to measure both aspects simultaneously, moral orientations and moral competence, producing two distinct set of scores for each aspect.

A competence test is defined by the task it contains. Thus a test of moral competence must contain a moral task: "In studying moral behavior we are concerned with studying action in which the subject gives up something or takes risks where not doing so would appear to be to his or her immediate advantage. [...] Thus, it is the overcoming of these situational pressures on either a verbal or a physical level that constitutes the test of moral behavior" (Kohlberg, 1984, p. 522).

As in Kohlberg's clinical moral judgment interviews, the MJT confronts the participant with a short story about a person in a dilemma situation. In the standard version of the MJT, two stories are used, the Doctor story and the Workers story. The participant is to give her or

his opinion on whether the actor's solution of the dilemma was wrong or right on a scale from -3 to +3. This part of the MJT is not used for assigning a test score. Rather it sets the stage for the moral task the participant is to solve.¹² Subsequently, the participants are asked to judge for each story six arguments supporting their opinion on the protagonist's decision, and six opposing their opinion on scales from -4 ("I strongly reject") to +4 ("I strongly agree"). The arguments have been selected (or constructed) to differ clearly in regard to their *moral quality*. To achieve this each argument represents one of the six types of moral orientation as defined by Kohlberg. Thus the six Kohlbergian types of moral orientation are represented by four arguments in the MJT.

The moral task of the MJT is to rate arguments supporting the participant's opinion on the protagonist's decision, and arguments opposing his or her opinion according to their moral quality but not in regard to their opinion agreement. For people at the lowest level of moral judgment competence, even the requirement to deal with arguments at all represents a very difficult task; they refuse to rate any of the arguments. A participant asked: "Why do I have to answer these questions after I did say my opinion on the issue". Participants at a somewhat more advanced level strongly agree with all arguments that support with their opinion, and strongly reject all arguments challenging their opinion. They find it difficult to dissociate themselves from bad, yet supportive arguments, and even more difficult to evaluate other people's arguments on the basis of their moral quality which oppose their opinion. Thus the pattern of responses to this moral task makes visible the *moral judgment competence* of the participants. As this competence develops, people start to rate arguments more and more in regard to their moral quality rather than in regard to their opinion-agreement, and the whole pattern becomes more *consistent in regard* to moral orientations and more *differentiated in regard* to one's own opinion.¹³

This competence aspect of moral behavior is indeed with the C-score, whereby the "C" stands for competence and cognitive aspect (for more details, see Lind, 2008 a). The C-score is calculated by analyzing the proportion of variance of an *individual's response pattern* that is determined by the moral quality of the arguments rather than by their opinion agreement, their context, or by any of the possible combinations of these three design factors of the MJT

¹² The common distinction between so-called 'preference-tests' and 'production-tests' does not apply here. Decisive is the distinction between moral *attitude* tests, which contain no task, and moral *competence* tests, like the MJT, which contain a difficult task, and cannot be faked upward (cf. Lind 2002).

¹³ Note that there is no "consistency" or "differentiation" of behavior per se, but both attributes must

(Lind, 2008 a). This proportion of variance is then multiplied by 100 to yield a score ranging from 0 to 100. The affective aspect (the profile of preferences for the six types of moral orientations) is simply measured, like in attitude measurement, by averaging the participants' evaluations of the arguments representing each moral orientation. These scores can range from -4 (strongly reject) to +4 (strongly accept).

For assessing the theoretical validity of test construction we have used two strategies. First we asked several Kohlberg-experts to rate the arguments' on Kohlberg's stage scale, and revised the arguments accordingly. Note that the arguments have not been submitted to some kind of empirical item selection in order to maximize their fit with statistical criteria. Second we used three well-corroborated theoretical predictions as validity criteria, namely the predictions a) that the preferences for the six Kohlberg-types of moral orientations are clearly ordered from high acceptance of Stage-6 moral reasoning to high rejection of Stage-1 moral reasoning (cf. Rest, 1969), b) that the six types of moral orientations are correlated in a way that neighboring types are more highly correlated than more distant stages, so the pattern of correlations form a 'quasi-simplex' (Kohlberg, 1958), and c) that moral orientations and moral competences are 'parallel', that is, that they are highly correlated. Obviously, these criteria are more clearly related to psychological theories, and more rigorous than traditional criteria of psychometrics. The MJT fulfills them all very clearly.

To test the hypothesis of affective-cognitive parallelism, we will use two methods. In the first analysis we divided the sample into nine sub-samples according to their C-score range. In the first group are all participants with a C-score between 0 and 9, in the second group all with a C-score between 10 and 19, and so on. Then we looked at each group's profile of preferences for the six moral orientations typical for Kohlberg's Stages. The parallelism hypothesis implies a) that the higher the participants' moral competence (C-score), the more clearly they prefer higher stage reasoning, and the more they reject lower stage reasoning. For obtaining estimates for effect size we looked at the moral preference ratings of groups with various level of moral judgment competence, analyzing polynomial contrasts for the preference profiles, and then converting the F -values to the effect size index r by the following formula, whereas df_j designates the number of categories minus 1, and df_i the number of cases minus 1, and r_{xy} the (nonlinear) correlation coefficient (Cooper & Hedges, 1994).

$$r_{xy} = \sqrt{\frac{df_j * F}{df_j * F + df_i}}$$

Another way to formulate this hypothesis is that the preference for of each Kohlbergian Stage of moral reasoning (as reflected by the judgments of the corresponding arguments) correlates in a predictable way with the participants' moral competence score: preferences for high stage reasoning should show a high positive correlation and this correlation should become lower and lower for preferences for lower stage reasoning. For preferences for the lowest stage of moral reasoning the correlation should be also high, but negative. To test this version of the parallelism hypothesis, we will therefore look at the profiles of six correlation coefficients.

It should be noted that the *a priori* probability of such a combined prediction is very small and, therefore, the falsifiability (Popper, 1968) and information value of this hypothesis is very high. There can be 720 possible outcomes, because there are six stages and the correlations with them can be ordered in $7! = 720$ different ways. Thus, the probability of a predicted order of correlations is 1 divided by 720, that is, $p = 0.0014$. Because this hypothesis is formulated as a universal prediction, the probability of accidental corroboration is extremely small and, therefore, its information value is extremely high.

The analysis of the parallelism hypothesis is based on MJT studies in Germany involving university students, apprentices, and prison inmates.

Findings

We have argued that Piaget's hypothesis implies that the cognitive aspect (represented by the C-index) and the affective aspect (represented by six attitude scales) correlate such that the higher the moral competence, the more lower stages of moral reasoning are rejected and the more higher (post-conventional) stages are accepted.

Early findings in Germany, where the first studies were done with the MJT, fully support Piaget's parallelism theory. *Figure 1* shows that, while all participants prefer higher to lower

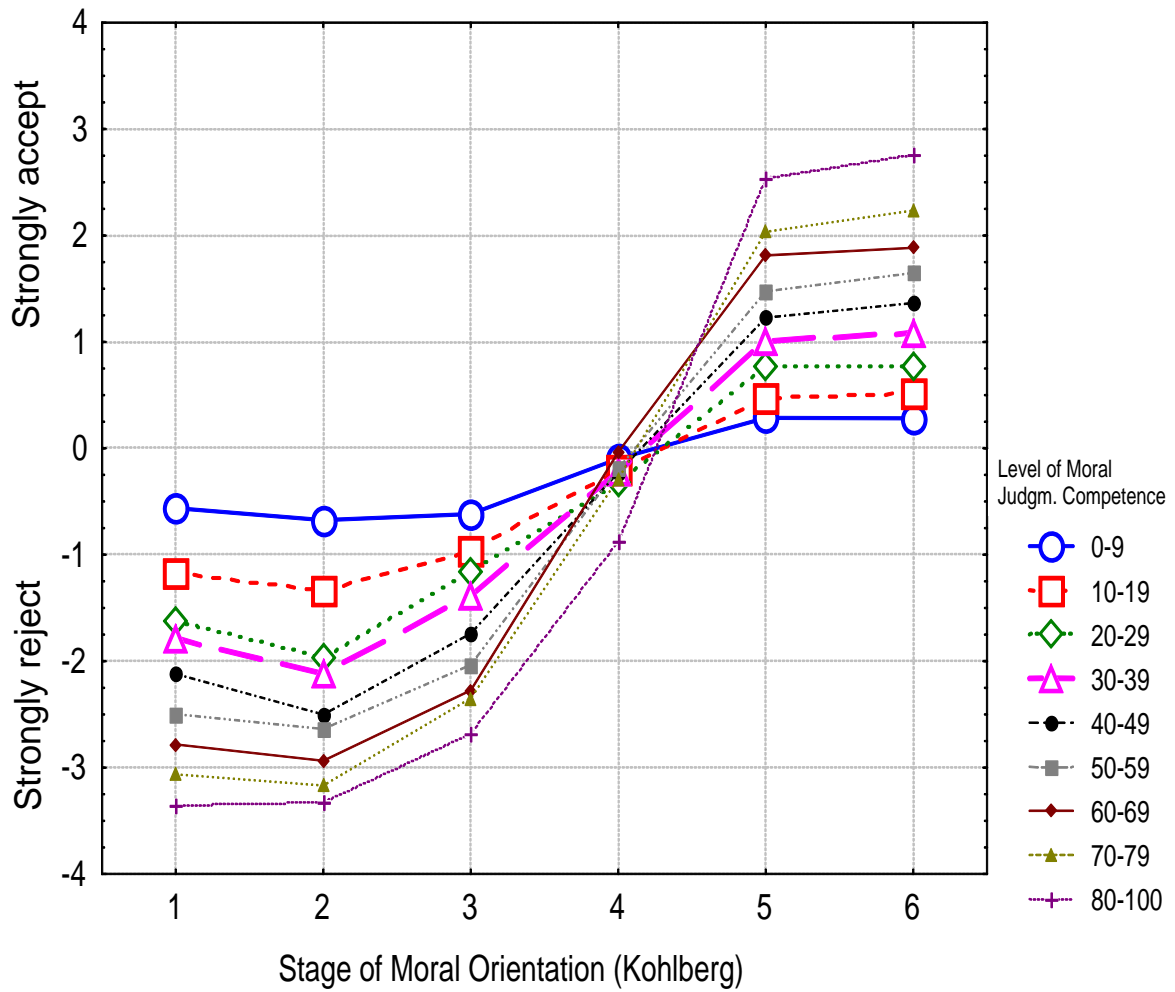
stages of moral orientations as cognitive-developmental theory predicts (Kohlberg, 1984; Rest, 1969), their *preference* for the *higher* stages (and the *rejection* of the lower stages) of moral orientation is stronger the higher their moral judgment competence is. It should also be noted that preference and rejection are graded as a direct function of the stage of orientation, and that the relationship is linear and without any exception. This is an unusually strong support for an unusually risky hypothesis, and we can regard Piaget's parallelism hypothesis as a well-founded cornerstone of moral development research.

This judgment is further corroborated by many findings with very different samples in Germany and in many other countries. Studying university students and juvenile delinquents also revealed the same pattern of correlations as predicted from the theory (Lind, 2002). As predicted, the correlations are very marked and their gradation is fully in line with the hypothesis. They are mostly below $r = -0.50$ for the lowest stage, and above $r = +0.50$ for the highest stage, and of intermediate size for the intermediate stages (*Figure 2*).

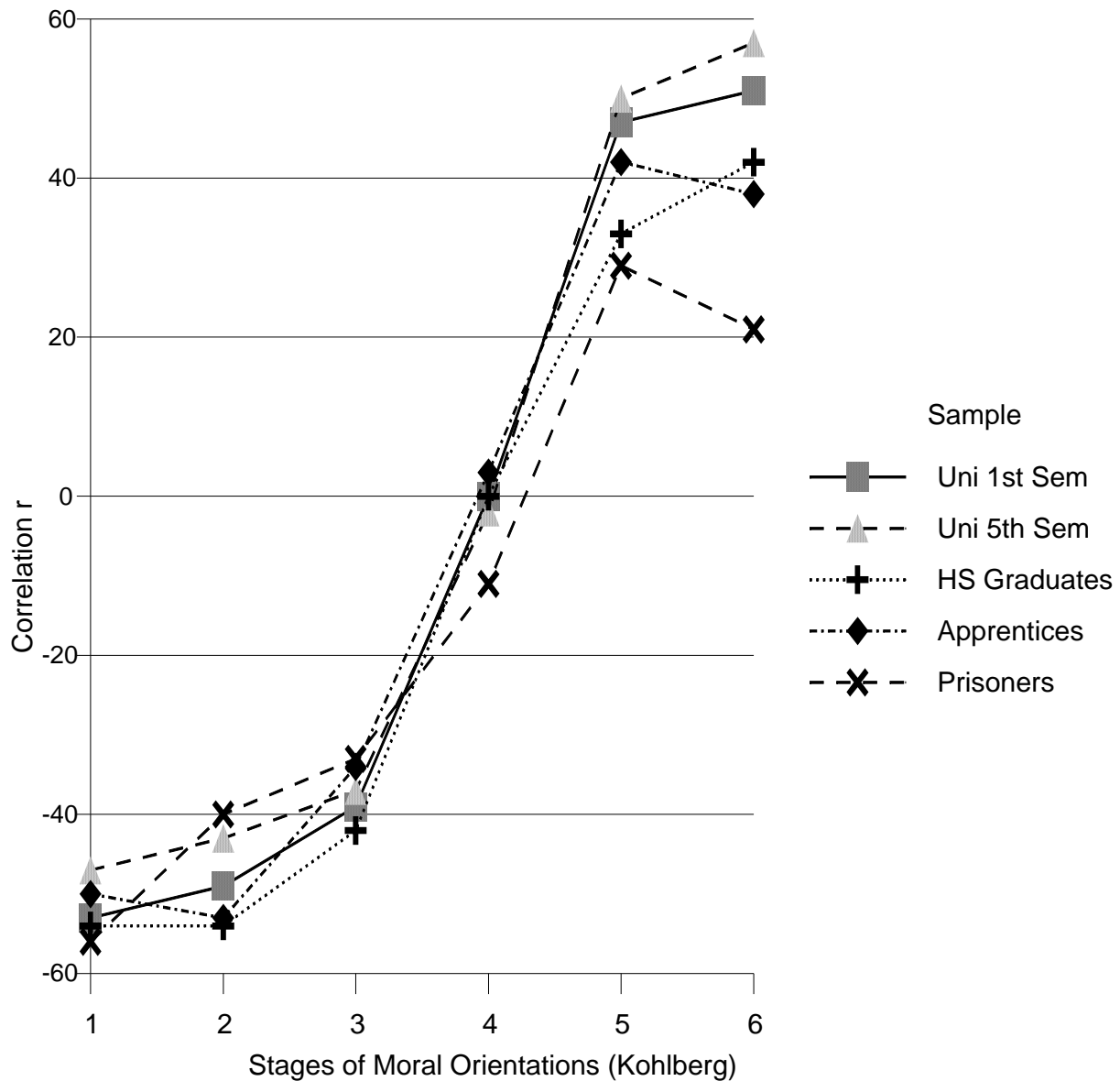
Until now the MJT has been translated into 39 languages and most have been certified as a cross-culturally valid measure of moral competence.¹⁴ Thus the hypothesis of affective-cognitive parallelism could be tested in many different countries and cultures. It was clearly supported – without exception – in *all* studies that I am aware of. For example, studies of university students in five other European countries (Austria, Germany, the Netherlands, Poland and Yugoslavia) revealed the same, invariant affective-cognitive parallelism (Lind, 2002).¹⁵ Because of the apparent universality of affective-cognitive parallelism, this phenomenon is now used as one of three validation criteria for new sub-tests for the MJT as well as for validating translations of the MJT. All new versions of the MJT must fulfill this criterion (and two more criteria) in order to be certified as valid, which ensures that all versions are not only semantically but also pragmatically equivalent (Lind, 2008 a). It must be noted that invalid items were revised to maximize their theoretical and inter-cultural validity but not to maximize the correlation of the C-score with age, or the difficulty of the MJT.

¹⁴ See http://www.uni-konstanz.de/ag-moral/mut/mjt-certification.htm#certified_versions

¹⁵ The findings from many more countries can be found at the above link. I wish to thank all authors of these studies for their permission to use their data for testing Piaget's parallelism hypothesis.



1 *Correlational parallelism* between moral affect and cognition: Acceptance of each of the six Kohlbergian types of moral orientations as a function of participants' moral judgment competence (C-score), MJT (index of moral judgment competence), in a sample of German first semester university students. Source: FORM-Project, 1977-1984 (see Lind, 2002). Interaction effect of Orientation and Competence-Level: $F(40,9830) = 92,43; p < 0,000; N = 2098;$ relative effect size $r = 0,52$.



2 *Correlational parallelism* between moral affect and cognition in five different samples: Profiles of correlations between participants' preferences of the six Kohlbergian types of moral orientation on the one hand (affective aspect) and the C-score (MJT) on the other (cognitive aspect), in samples of 1st semester university students (N = 2098), 5th semester university students (N = 812; both FORM-project), high school graduates (N = 516; also FORM; Lind, 1978), Swiss apprentices (N = 579; HASMU-project by Fritz Oser and his colleagues, reported in Lind, 2002), and juvenile prisoners (N = 58; Wischka, 1982).

Implications for Moral Psychology and Education

Clearly, Piaget's paradigm of distinct-yet-not-separate aspects sheds light in the different ways how we organize our educational systems and how we design the school curriculum regarding moral education and more. In the past, the separation of cognition and affect as different *substances* or *components* had a great impact on the organization of our educational system, curriculum construction and educational assessment and evaluation. The organization of our educational system reflects the component model and leaves little room for an integrative approach to moral education as implied by the aspect model. Each component, it seems, "has" its own teaching subject and department of education. Everyone trying to implement an integrative, affective-cognitive approach of moral education in schools, can tell painful stories about the misfit of our institutions of education for such an approach. If moral education is to be effective, we must acknowledge that morality is as much a cognitive competence as it is an affective disposition, and that it is moral competence that must be fostered, not moral orientations. Our studies let us conclude that fostering moral judgment competence also strengthens principled moral reasoning and behavior. The more it is developed, the more clearly adolescents (and adults) discern the inadequacy of low stage reasoning, and the more strongly they adhere to moral principles in their reflection on moral decisions. Secondly, ability to apply moral principles to one's judgment behavior also leads to better decision-making. Participants with high ability clearly make a decision (in one or the other direction) yet they refrain mostly from taking too extreme stances on an issue, whereas participants with low judgment competence tend either to take an extreme stance (in either direction) or no stance at all.

As Piaget has argued, and our research confirmed, without affective and emotional arousal there is little learning and hardly a lasting effect of learning. And without taking the cognitive aspects of moral affects into account, there is no moral development from the level of black-and-white moral thinking (which is associated with a high probability to resort to violence as a means to "promote" the good) to the level of more integrated and differentiated moral judgment facilitating nonviolent ways of conflict resolution like moral discourse, mediation and peaceful negotiation. In other words, constructivist moral education based on Piaget's parallelism theory of moral behavior and development eventually strengthens students' decision-making capacity without pushing them into one direction or the other (as

indoctrination would do). Thus fostering students' moral competence agrees well with the moral principles of a democratic way of life. It strengthens students' ability to speak up and listen to others, and their ability to participate in a democratic discourse and non-violent conflict resolution (Lind, 2008 b).

Conclusions

Piaget's concepts and hypotheses have laid the ground for a new paradigm of moral psychological research and educational practice, which, it seems, has not yet been taken full advantage of. Piaget's aspect concept overcomes the problems implied by the component model, rooted in the conceptual realism of Plato and Descartes. It has helped to design new methods of measurements which allow simultaneous assessment of cognitive and affective aspects of moral behavior (and of other behaviors as well, of course).

Methodologically, the present study supports both the fruitfulness of Piaget's aspect model and the empirical validity of his parallelism hypothesis. Neither a purely 'cognitive' nor a purely 'affective' approach to the measurement of moral judgment behavior is warranted. Neither aspect can be adequately assessed without reference to the other. There is no *pure structure* of human behavior irrespective of content or direction and energy. When we talk about behavioral consistency, we always have to define consistency (or inconsistency) *in regard* to some behavioral standard, norm or principle. In contrast to chemistry, in psychology there is no *consistency per se*. Only when we define consistency in regard to some moral orientation, we can distinguish principled judgment from rigid judgment, nor could we tell apart situated judgments from amoral, erratic judgments (Eyferth, 1959). Neither is there *pure affect* or attitude irrespective of the cognitive processing of the situation which triggers the behavior. Assessing both aspects simultaneously allows us a) to distinguish moral consistency from rigidity of opinion, and b) it also lets us determine if a participant's pattern of behavior exhibits clearly structured moral orientations, or no moral orientation at all, or a highly differentiated moral judgment.

In many studies affect and cognition are misconceived of as separable components (Rest 1984; Rest et al., 1999; see also Beck, 1995, p. 117; Gibbs & Schnell, 1985, p. 1078) or are even placed in separate *domains* of educational objectives (Krathwohl et al., 1964; Tomlinson-Keasey & Eisert, 1981). In some studies, moral cognition and affect have been not only

separated (which is not possible, as we have seen) but have even been opposed to each other with the question as to which is the more important or more *real* component (Emler et al., 1983; Hogan & Emler, 1995; Haidt, 2001; Greene & Haidt, 2002; Zajonc, 1980). Within the dual aspect paradigm, this question is meaningless. Interestingly, the authors of these studies not only place morality in the affective domain and confine its assessment to attitude measurement (Emler et al., 1983; Zajonc, 1980) and to neurological imaging of emotional processes (Haidt, 2001; Greene & Haidt, 2002), but they also show blatant disregard for the methodological paradigm shift triggered by Piaget's aspect model. As Scott (1968) and Lind (2010 c) have shown, this belief has led to a stagnation of attitude research. The negative consequences of the component model for educational research and educational practice have been pointed out by Sprinthall et al. (1994) who noted that Bloom et al.'s taxonomy of educational objectives artificially separates affect and cognition into different domains of behavior and so undermines an integrative approach to teaching and learning. "In the separation between the social and the cognitive," the former AERA president Alan Schoenfeld (1999) noted, "some fundamentally important issues such as affect and motivation have fallen between the cracks. We need to build new frameworks and perspectives that do justice to all of these. And we need new methods to inform the work done within those perspectives" (p. 5).

The Dual-Aspect-Model suggested by Piaget proved to be a real paradigm shift in psychological research. It made it possible to create a new experimentally designed instrument, the *Moral Judgment Test*, that allows us to measure the two aspects of moral behavior simultaneously without separating them as components. This, in turn, allowed us to test Piaget's hypothesis of affective-cognitive parallelism adequately. Our findings show with great clarity that there is indeed a strong parallelism: *People prefer higher moral orientations, and reject lower moral orientations, the higher their ability to make moral judgments*, that is, judgment based on moral orientations rather than on opinion-agreement. However, we assume that moral orientations and moral competence are only parallel in normal situations. This parallelism may break down if the testing situation involves 'high stakes' for the tested person. If participants sense that their answers to the test will trigger gratifications or negative sanctions, they will try to do what they believe the test administrator (or his commissioner) expects of them. For example, if the test of moral judgment behavior is used for deciding on the admission to an educational program the participants will have a strong incentive to fake the scores "upward" (Lind, 2002). The dissolution of cognitive-affective parallelism becomes

evident in the classical study by Emler et al. (1983), in which participants are instructed to simulate the moral preferences of other people. This study demonstrated that people can simulate almost any moral preference, while we have strong evidence that they cannot simulate other people's moral judgment competence, if this is higher than their own (Lind, 2002). Affective-cognitive parallelism may also break down when moral development regresses. As our research shows, moral competence can regress when there are no opportunities to practice it, however, moral orientations do not. If the support through schooling ceases when students have not yet reached a critical level of moral development, the ability to apply these orientations in everyday life can erode dramatically.¹⁶

References

- Anastasi, A. (1958). *Differential Psychology*. New York: McGraw-Hill.
- Anderson, N. H. (1991). Moral-social development. In: N. H. Anderson, *Information Integration Theory, Volume III: Developmental*, pp. 137 - 187. Hillsdale, NJ: L. Erlbaum.
- Beck, K. (1995). Aspekte der moralischen Urteilsbildung bei kaufmännischen Lehrlingen – Methodologische Probleme und empirische Befunde [Aspects of moral judgment development in merchandise apprentices]. In: E. H. Hoff & L. Lappe, eds., *Verantwortung im Arbeitsleben*, pp. 114 - 128. Heidelberg: Roland Asanger Verlag.
- Colby, A. (2008). Fostering the moral and civic development of college students. In: L. P. Nucci & D. Narvaez, eds., *Handbook of moral and character education*, pp. 391-413. New York: Routledge.
- Colby, A. & Kohlberg, L. (1984). Invariant sequence and internal consistency in moral judgment stages. In W. M. Kurtines & J. L. Gewirtz, eds., *Morality, moral behavior, and moral development*, pp. 41 - 51. New York: John Wiley & Sons.
- Cooper, H., & Hedges, L. V., eds. (1994). *The handbook of research synthesis*. New York: Russell Sage.

¹⁶ In the case of *pseudo*-regression, the opposite phenomenon may occur. Pseudo-regression (or the *Raskolnikov syndrom* as Kohlberg & Kramer (1969) named it after the character in Dostojewski's novel *Crime and Punishment*) means that people start to prefer lower moral orientations than they have acquired but retain their judgment competencies. This has been observed in adolescents during their transition from high school to college, when they free themselves from the normative context of family life (see Lind, 2010 a).

- DeVries, R. (1997). Piaget's social theory. *Educational Researcher*, 26(2), 4 - 17.
- Emler, N., Renwick, S. & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology*, 45, 1073 - 1080.
- Eyferth, K. (1959). Integration und Starrheit [Integration and rigidity]. *Psychologische Rundschau*, 9, 159 - 169.
- Galton, F. (1889). *Natural inheritance*. London: Macmillan.
- Gibbs, J. C., & Schnell, S. (1985). Moral development 'versus' socialization: A critique. *American Psychologist*, 40, 1071-80.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96, 5, 1029 – 1046.
- Greene, J. D. (2008). The secret joke of Kant's soul. In: W. Sinnott-Armstrong, ed., *Moral psychology. Volume 3: The neuroscience of morality: Emotions, brain disorders, and development*, pp. 39 - 79. Cambridge, MA: The MIT Press.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 4, 814 - 834.
- Hartshorne, H. & May, M. A. (1928). *Studies in the nature of character. Vol. I: Studies in deceit, Book one and two*. New York: Macmillan.
- Helkama, K., Uutela, A., Pohjanheimo, E., Salminen, S., Koponen, A., & Rantanen-Vantsi, L. (2003). Moral reasoning and values in medical school: a longitudinal study in Finland. *Scandinavian Journal of Educational Research*, 47, 4, 399 - 411.
- Higgins, A. (1995). Educating for justice and community: Lawrence Kohlberg's vision of moral education. In: W. M. Kurtines & J. L. Gewirtz, eds., *Moral development: An introduction*, pp. 49 - 81. Boston: Allyn & Bacon.
- Hogan, R. & Emler, N. (1995). Personality and moral development. In W. M. Kurtines & J. L. Gewirtz, eds., *Moral development. An introduction*, pp. 209 - 227. Boston: Allyn and Bacon.
- Hopkins, K. D. & Glass, G. V. (1978). *Basic statistics for the behavioral sciences*. Englewood, NJ: Prentice-Hall.
- Kohlberg, L. (1964). Development of moral character and moral ideology. In: M.L. Hoffman & L.W. Hoffman, eds., *Review of Child Development Research*, Vol. I, pp. 381 - 431. New York: Russel Sage Foundation.
- Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach.

- In T. Lickona, ed., *Moral development and behavior: Theory, research and social issues*, pp. 31 - 53. New York: Holt, Rinehart & Winston.
- Kohlberg, L. (1984). The meaning and measurement of moral judgment. In: L. Kohlberg, *Essays on moral development, Vol. II, The psychology of moral development*, pp. 395 - 425. San Francisco, CA: Harper & Row.
- Kohlberg, L. & Kramer, R. (1969). Continuities and discontinuities in childhood and adult moral development. *Human Development*, 12, 93 - 120.
- Krathwohl, D. R., Bloom, B. S. & Masia, B. B. (1964). *Taxonomy of educational objectives. Handbook II: Affective domain*. New York: McKay.
- Kuhn, D., Langer, J., Kohlberg, L. & Haan, N. (1977). The development of formal operations in logical and moral judgments. *Genetic Psychological Monographs*, 95, 97 - 188.
- Lind, G. (1978). Wie misst man moralisches Urteil? Probleme und alternative Möglichkeiten der Messung eines komplexen Konstrukts [How can we measure moral judgment? Problems and alternative possibilities if measuring a complex construct.]. In G. Portele, ed., *Sozialisation und Moral*, pp. 171 - 201. Weinheim: Beltz.
- Lind, G. (1989). Essay review: 'The measurement of moral judgment' by Anne Colby, Lawrence Kohlberg et al.. *Human Development*, 32, 388 - 397.
- Lind, G. (2000). Moral regression in medical students and their learning environment. *Revista Brasileira de Educacao Médica*, 24(3), 24 - 33.
- Lind, G. (2002). *Ist Moral lehrbar? Ergebnisse der modernen moralpsychologischen Forschung*. [Can morality be taught? Research findings from modern Moral Psychology.] Berlin: Logos-Verlag. [English publication is in preparation.]
- Lind, G. (2008 a). The meaning and measurement of moral judgment competence revisited – A dual-aspect model. In: D. Fasko & W. Willis, eds., *Contemporary philosophical and psychological perspectives on moral development and education*, pp. 185 - 220. Cresskill, NJ: Hampton Press.
- Lind, G. (2008 b). Teaching students to speak up and to listen to others: Cultivating moral democratic competencies. In: D. E. Lund & P. R. Carr, eds., *Doing democracy and social justice in education: Political literacy for all students*, pp. 319 - 335. New York: Peter Lang Publishing.
- Lind, G. (2009). *Moral ist lehrbar. Handbuch zur Theorie und Praxis moralischer und demokratischer Bildung*. [Morality can be taught. Handbook of theory and practice of moral and democratic education]. München: Oldenbourg.

- Lind, G. (2010 a). Growth and regression in cognitive-moral development of young university students. In: C. G. Harding, ed., *Moral Dilemmas and Ethical Reasoning*. 2nd edition, pp. 99 - 114. New Brunswick, NJ: Transaction Publisher.
- Lind, G. (2010 b). The theory of moral-cognitive development: a socio-psychological assessment. In G. Lind, H. A. Hartmann, & R. Wakenhut, eds., *Moral Judgment and Social Education*, pp. 21 - 53. Rutgers, NJ: Transaction Books, 2nd edition.
- Lind, G. (2010 c). Attitude change or cognitive-moral development? How to conceive of socialization at the university. In G. Lind, H. A. Hartmann & R. Wakenhut, eds., *Moral Judgment and Social Education*, pp. 173 - 192. Rutgers, NJ: Transaction Books, 2nd edition.
- Loevinger, J. (1976). *Ego development*. San Francisco: Jossey-Bass.
- May, M.A. & Hartshorne, H. (1926). First steps toward a scale for measuring attitudes. *The Journal of Educational Psychology*, 17, 145 - 162.
- Meehl, P.E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft Psychology. *Journal of Consulting and Clinical Psychology*, 46, 806 - 834.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156 - 166.
- Miller, G.A. (1969). *Psychology, the science of mental life*. Harmondsworth: Penguin (Original 1962).
- Montada, L. (1993). Moralische Gefühle [Moral feelings]. In W. Edelstein, G. Nunner-Winkler & G. Noam, eds., *Moral und Person*, pp. 259-303. Frankfurt: Suhrkamp.
- Nunner-Winkler, G. (1989). Verantwortlichkeit für andere. Das Problem der 'positiven' Pflichten [Responsibility for others. The problem of 'positive' duties]. In: G. Lind & G. Pollitt-Gerlach, eds., *Moral in 'unmoralischer' Zeit*, pp. 33 - 54. Heidelberg: Asanger.
- Oliveira-Souza, R.de, Ignácio, F. A., & Moll, J. (2008). The antisocials amid us. In: W. Sinnott-Armstrong, ed., *Moral psychology, volume 3: The neuroscience of morality: emotion, brains disorders, and development*, pp. 151 - 171. Cambridge, MA: The MIT Press.
- Piaget, J. (1951). *Play, dreams and imitation in childhood*. New York: Norton.
- Piaget, J. (1965). *The moral judgment of the child*. New York: The Free Press.
- Piaget, J. & Inhelder, B. (1969). *The psychology of the child*. Basic Books.
- Piaget, J. (1971). The theory of stages in cognitive development. In D. R. Green, ed., *Measurement and Piaget*, pp. 1-11. New York: MacGraw-Hill.

- Piaget, J. (1976). The affective unconscious and the cognitive unconscious. In: B. Inhelder & H. H. Chipman, eds., *Piaget and his school*. New York: Springer, pp. 63 - 71.
- Piaget, J. (1981). *Intelligence and affectivity: Their relation during child development*. Palo Alto, CA: Annual Reviews. (Originally published 1954).
- Popper, K. R. (1968). *The logic of scientific discovery*. London: Hutchinson.
- Prehn, K., Wartenburger, I., Mériaux, K., Scheibele, C., Goodenough, O., Villringer, A., van der Meer, E. & Heekeren, H. (2008). Influence of individual differences in moral judgment competence on neural correlates of socio-normative judgments. *Social Cognitive and Affective Neuroscience*, 3(1), 33-46.
- Rest, J. R. (1969). Level of moral development as a determinant of preference and comprehension of moral judgments made by others. *Journal of Personality*, 37(1), 220 - 228.
- Rest, J. R. (1984). The major components of morality. In: W. M. Kurtines & J. L. Gewirtz, eds., *Morality, moral behavior, and moral development*, pp. 24 - 38. New York: Wiley.
- Rest, J. R. (1988). *Moral judgment: an interesting variable for higher education research*. ERIC, August, HE 021 280.
- Rest, J. R., Narvaez, D., Bebeau, M.J. & Thoma, S.J. (1999). *Postconventional moral thinking. A Neo-Kohlbergian approach*. Mahwah, NJ: Erlbaum.
- Schoenfeld, A. H. (1999). Looking toward the 21st century: Challenges of educational theory and practice. *Educational Researcher*, 28, 4 - 14.
- Scott, W.A. (1968). Attitude measurement. In: G. Lindzey & E. Aronson, eds., *Handbook of social psychology*, pp. 205 - 270. Reading MA: Addison-Wesley.
- Sprinthall, N. A., Sprinthall, R. C. & Oja, S. N. (1994). *Educational psychology. A developmental approach*. 6th edition. New York: McGraw-Hill.
- Tomlinson-Keasey, C. & Eisert, D. C. (1981). From a 'Structure d'Ensemble' to separate organizations for cognitive and affective development. *Human Development*, 5, 1 - 19.
- Wilson, M. (2005). *Constructing measures. An item response modeling approach*. Mahwah, NJ: Erlbaum Publishing Inc.
- Wischka, B. (1982). *Moralisches Bewußtsein und Empathie bei Strafgefangenen*. University of Marburg: Unpublished diploma thesis.
- Wittgenstein, L. (1953). *Philosophical investigations*. Translated by G. E. M. Anscombe. New York: The Macmillan company.
- Wuttke, J. (2007). Uncertainties and bias in PISA. S. T. Hopmann, G. Brinek & M. Retzl,

eds., *PISA zufolge PISA*, pp. 241 - 264. Berlin: LIT-Verlag.

Zajonc, R. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151 - 175.