

The Art of Experimental Psychology¹

Ideas for Studying Moral Competence and Beyond

Georg Lind²
Konstanz, Germany

Why?

I am often asked for ideas for studies on the nature, relevance and teachability of moral competence. Some ask me about my *Moral Competence Test*, (Lind 2016) others about research findings, still others about a good research question which would be worthwhile to study. Quite frequently I also get outlines, drafts, or final reports for commenting. During the more than 40 years of research and teaching in this area, I have also read many scientific articles to prepare my own research. Much of what I am reading is stimulating, well-done, and well-reported. But unfortunately much also suffers from serious deficiencies: e.g., no research question at all (“explorative research”); the research object ill-defined or not defined at all; the measurement instruments do not match the author’s intentions; statistics like “significance” are not understood and wrongly applied; statistical instead of psychological models are used as the basis of the research; “transformed” and “normalized” data, instead of raw data and mean scores, are reported which makes it impossible to compare the findings across studies; significance tests instead of (absolute) effect sizes are reported; coefficients of correlations and regression are taken as *prove* for causal relationships.

Yes, “authors” must be held responsible for their texts. But it is not only their fault. The methodological education of psychologists still has not improved much since fifty years ago when I studied psychology. Especially statistic textbooks have not become any better. Most do not explain that “significance” does not mean importance or effect but merely means that a sample is large enough for detecting small differences (Shaver 1993). They do not explain that “factors” and “paths” are only names for mathematical formulas but have no psychological meaning. They do not explain that “normal” distributions of data are not normal at all; they can only be found where pure chance is at work like roulette and pure random error of observations. (Bortz 1983) Most human traits are not “normally distributed.” (Micceri 1989) Significance testing has impeded the advancement of psychology. (Carver 1993; Cohen 1994)

The biggest deficiency is the prevailing “inductivist” view of how science works: Inductivists believe that psychological theories emerge from observation and measurement data (“evidence”). But this is wrong. Theories are not inherent in data. Rather we invent them in our heads. If we want psychology to grow we must adopt a “deductivist,” experimental view: we must submit our theories to rigorous empirical and experimental tests. Only well designed experiments provide data which can weed out wrong theories and confirm valid theories. Theories can be found in scientific literature as well as in every-day. All our decision-making is guided by – tacit or explicit – theories which ask for verification. Some examples: “Without religion there can be no morality.” “Moral values are imprinted on the individual by

¹ Last revision: February 4, 2019.

² Contact: Dr. Georg Lind (apl. Prof. em., Konstanz), Email: georg.lind@uni-konstanz.de.

society.” “Moral behavior is determined by the values a person holds.” “Moral competence is most relevant for behavior.” “Moral competence can be fostered through good education.” Which of these theories are valid? Which can be trusted?

I have found and tested the ideas which I pass on here during the past forty years as a researcher, teacher and citizen. You may use them not only for studies on moral psychology but also in fields of psychological, educational and social science research. If they do not seem clear and consistent to you, please correct them and let me know. If you feel that they are too short-hand, please look at the references listed below and at my other methodological writings: <http://www.uni-konstanz.de/ag-moral/b-liste.htm>

If these ideas or guidelines make you aware of the problems of psychological research, and stimulate your own ideas how to make good psychological research, it was already worth writing them up.

Good research starts with a good research question

A good question can be easily recognized. If your findings let you answer a question with a clear “yes” or “no,” or “yes, to a [quantifiable] extent” or “yes, if an [identifiable] condition is given,” then you have asked a good question. For example, if you can answer the question “Does moral competence reduce criminal behavior?” with “yes” or “no,” your findings advance our knowledge about moral competence regardless whether its findings confirm a hypothesis or not. If findings are positive, the study reinforces endeavors to foster moral competence. If they disconfirm it, the study provides arguments to search for other traits which lower criminal behavior.

In contrast typical statistical research questions like “Are there any differences between males and females in regard to that moral competence?” have neither a scientific nor a practical value. Actually one can spare research. This hypothesis is always true if the sample is large enough. Similarly research questions like “Does trait A correlate with trait B?” lack psychological meaning. Statistical correlation does not tell us whether “A influences B or B influences A.” Correlation does not even prove a relationship. There can be spurious correlations and spurious non-correlations.

Spurious correlation: For example, a correlation between moral competence and age can mean (a) that age causes moral development, or (b) that actually education causes moral development – but because education is correlated with age, age is also statistically correlated with moral development although there is no psychological relationship.

Spurious non-correlation: If you do not find a (linear) correlation, the reason could be that the correlation is non-linear, or asymmetric, or that there is no variation of the variables in your sample. For mathematical reasons the mostly used Pearson coefficient is always zero if the variance in one variable is zero. Below I will show better ways to measure effect size and relationships (which is actually the same).

The bottom line: Statistical models tell us little, if anything, about the nature, relevance and teachability of psychological traits. Only psychological models do.

When, in the early 1970ies, I looked at different psychological models and concepts which could be a good starting point for my own research and could be a basis for contributing something valuable to the world, I did not find many. In fact I found only one field which was genuinely psychological, which was pursuing new methods of psychological research, and which seemed to contribute to democracy as a way of living together: the cognitive-developmental theories of moral behavior by Jean Piaget (1965) and Lawrence Kohlberg (1956; 1985). Both scholars were far ahead of their time and are still far ahead of many psycho-

logists who do not understand the significance of their theories. I felt that their theories were in some respects also ahead of their own research methods. They had well grounded theories but had no objective research methods for studying their validity. (Lind 1989) So I made it my task to close this gap. (Lind 1978) I developed a valid and objective measurement of moral competence, the *Moral Competence Test*. (MCT, formerly called MJT; Lind 2016) The MCT shows that many of Piaget and Kohlberg’s insights are testable and empirically true (see below), yet also that some parts of their theories (e.g., the postulate of non-regression of moral development) have been refuted by data. (Lind 2002)

The psychological research questions and hypotheses concerning moral competence can be sorted into three broad categories: about its nature, its relevance, and its teachability. Some have already been studied intensively while others still need to be studied.

There is hardly any other human trait like moral competence for which so many grounded and testable hypotheses about its *nature*:

(1) Hypotheses concerning the existence of moral orientations and moral competence. We can characterize people’s behavior by certain traits and that these traits really exist – if they manifest themselves through their behavior. If they are only statistical constructs (“latent variables,” “factors”), we do not need to deal with them in psychology.

(a) Hypothesis: Morality is real, but it is not a “thing” like a chair or a leg, instead it is an aspect or attribute of behavior. Morality is not a single aspect but it consists of two aspects, namely *moral orientations* (affective aspect) and *moral competence* (cognitive aspect). These aspects are clearly *distinguishable*, but *inseparable*. (Piaget 1976)

Finding: Piaget’s argument is supported by the following hypotheses and findings.

(b) Hypothesis: As Piaget (1976) asserts, that moral orientations and competence are preconscious but real because they manifest themselves in behavior, namely in the patterns of people’s responses to certain tasks and questions. (Lind 2016, ch. 4)

Finding: With the experimentally designed *Moral Competence Test* it is possible to measure both aspects distinctly and objectively as attributes of the response pattern of an individual person. (Lind 2016) This means that both aspects do not need to be assessed indirectly through clinical interviews or statistical correlations. On the basis of forty years of research it seems safe to say that moral orientations and competence can be measured objectively and validly. (Lind 1978; 1980; 2016) The two aspects and the two layers of the moral self are depicted in this graph:

The Dual-Aspect-Dual-Layer Model of the Moral Self

	Affective Aspect	Cognitive Aspect
The layer of conscious reasoning and judgment (Moral philosophy; ethics)	Ethical sentiments and principles as verbalized in a conversation	Ethical judgment and reasoning
The layer of <i>manifest</i> behavior and feelings (Moral psychology, biology)	Moral orientations as manifested in a <i>pattern of behavior</i>	Moral competence

* Adapted from: Georg Lind (2016). *How to Teach Morality...* Berlin: Logos.

(c) Hypothesis: As Piaget (1976) also asserts, moral orientations and moral competence cannot be separated like components (as the weight and the form of a ball cannot be separated from the ball like the components air and paint), but they are distinct aspects; they develop differently and they need to be treated differently in education.

Findings: In fact, as Socrates already observed, both aspects are distinct: all people want to be good but only a few have the capability to be as good as they want to be. Whereas moral orientations seem to be inborn and are similar in all people, moral competence differs widely among people. Whereas moral orientations hardly change, moral competence develops in a favorable learning environment.

(d) Hypothesis: Like all competences, moral *competence* cannot be faked upward, whereas moral orientations can be faked in any direction.

Findings: In two faking experiments using the *Moral Competence Test* (MCT), the participants could not fake their moral competence upward. (Lind 2002) In contrast participants could fake moral orientations in any direction. (Emler et al., 1983)

(e) Hypothesis: According to Kohlberg (1956; 1984) six different types of moral orientations can be distinguished, and form a hierarchy of adequacy: people prefer higher types of moral orientation as more adequate for solving a moral dilemma than lower ones.

Findings: Many studies have consistently shown that nearly all people, regardless of age, gender or culture, indeed put Kohlberg's six types of moral orientation into the same hierarchical preference order as Kohlberg has predicted. (Rest 1969; Lind 1978; 2002)

(f) Hypothesis: These six moral orientations are not only hierarchically ordered but are also empirically related in such a way that "neighboring" types are more similar and correlate more closely than more "distant" types.

Finding: This pattern of relationships has been consistently found in many MCT-studies. (Lind 1978; 2002)

(g) Hypothesis: Piaget (1965) assumed that the cognitive aspect of moral behavior (moral competence) and the affective aspect (moral orientation) are "parallel." One meaning of parallel is that both aspects correlate highly with each other: the higher the moral competence of people, the more clearly they reject inadequate (low) moral orientations, and the more clearly they accept adequate (high) moral orientations.

Finding: All MCT-studies have confirmed this hypothesis amazingly well. (Lind 1986)

(h) Hypothesis: Moral orientations hardly show any statistical correlations with behavior which are believed to have a moral basis. This is not because moral orientations are unimportant but because (the popular) linear correlations are always zero if the variance of one of the variables (moral orientation) is zero. In contrast, moral competence, which varies greatly, should correlate strongly with behavior.

Finding: (see below).

This clear confirmation of all hypotheses about the nature of the moral self in many studies in the past four decades by many different scholars, lets us conclude that moral orientations and moral competence are real and that they are distinct aspects of moral behavior.

The most central aspect is *moral competence*, because this develops, varies between people, is highly relevant for behavior, and should be fostered in education. It is defined as *the ability to solve problems and conflicts on the basis of moral orientations or principles, through*

deliberation and discussion, instead of through violence, deceit, submission to others, or ignoring. (Lind 2016)

The most important application of this knowledge is measurement. If our knowledge about the nature of moral competence and moral orientations is well confirmed, it can be used to check the validity of translations of the MCT. If the translation produces unexpected findings we doubt the translation. In every case in which studies had raised doubts about the validity of a translated version, the author confirmed this doubt and revised the test successfully. In a second validation study the anomalies of the translated test had disappeared.

(2) Research questions concerning the *relevance* of moral competence for behavior.

Since we know now that moral orientations and moral competence really exist, we can turn now to hypotheses stating that these two aspects are relevant for behavior. This relevance is widely believed by many people, but also often questioned. (Althof et al. 1999)

(a) Hypothesis: *Moral orientations* are essential for moral behavior by definition. A behavior is called moral only if it is (at least partly) motivated by internal moral principles and ideals.

(b) Hypothesis: Moral competence, we have seen, is defined as the ability to solve problems and conflicts on the basis of moral orientations or principles, through deliberation and discussion, instead of through violence, deceit, submission to others, or ignoring. Therefore, moral competence should show up in many behaviors: Honesty, staying out of troubles, keeping contracts, helping people in distress, learning, decision-making, and so on.

Findings: Many empirical and experimental studies have confirmed this hypothesis. (Kohlberg 1984; Hemmerling 2014; Lind 2016; see also the appendix for an overview and more references)

More research is needed:

(c) Many areas of behavior, in which moral competence may play an important role, have not yet been studied or not properly studied: e.g., its impact on environmental protection, on judging the trustworthiness of other people (like life-partners, business-partners, and politicians), on the quality of professional work, e.g., of teachers, physicians, politicians. Does moral competence impact the democratic quality of social institutions like family, neighborhoods, communities, nations and trans-national organizations (EU, UN).

(c) Moral competence may also have an indirect influence on behavior by impacting human traits which are important for social behavior, e.g., *pluralistic ignorance, stereotypes/prejudice, tolerance of ambiguity, ego-strength, and conformity*. Many people base their decisions on what they believe about the traits and behaviors of other people. If they believe, e.g., that other people would not cooperate or do help them when they needed their help, even though these would help (= pluralistic ignorance), they also do not cooperate or help – except if they could develop high moral competence. (DIT-study by Jacobs 1975) If people believe that *all* foreigners tend to use violence or deceit for solving problems and conflicts, even though they have no proof for their belief (= stereotype or prejudice), then they do not want to live in their neighborhood or their country. (Allport 1954) The ability to tolerate ambiguity and ego-strength also seem to be impacted by moral competence. (Lind et al. 2010) A small study by Mofakhami (2018) indicates that moral competence also influences *Ash*-conformity.

(c) Does the impact of moral competence depend on certain conditions (like age, social climate)? Do certain conditions prevent this trait from becoming relevant?

(d) Can one-time findings be replicated over time and across cultures? Replication studies are essential for the advancement of science. (Shaver 1993) Unfortunately, in psychology many journal editors and dissertation reviewers do not understand the importance of replications and reject their publication. This impedes science and needs to be changed.

In sum, so far all hypotheses regarding the relevance of moral competence for social behavior and for living together in a democracy have been clearly confirmed. No negative impact was found. This is rare in psychology. Thanks must go to Jean Piaget and Lawrence Kohlberg who have paved the way.

(3) Research questions concerning the teachability of moral competence.

(a) Hypothesis: Like all abilities, moral competence can be fostered most effectively by providing challenging opportunities or tasks for using it.

Findings: Many studies have shown that moral competence does not develop by itself, and cannot be increased by social force, but its development depends on (good) education (Lind 2002). Studies also show that education is “good” if it provides the learner with a minimum of opportunities of responsibility-taking and guided reflection. (Lind 2000; Schillinger 2006) And many intervention stories demonstrate that specific methods like Blatt-Kohlberg-method of dilemma discussion and the *Konstanz Method of Dilemma-Discussion* (KMDD) are even more effective and also very efficient: They produce strong effects with relatively little investment of time. But they are only effective if the teacher who uses these methods is well trained. (Lind 1989; 2002)

Specific conditions and side-effects need to be researched yet:

(a) Hypothesis: Certain conditions are required so that certain methods and programs can be effective?

Finding: The KMDD can foster moral competence effectively only if the teacher is well trained to use it. (Lind 2016)

(b) Hypothesis: Some conditions may prevent positive effects of good education and methods like the KMDD.

Findings: Students with dogmatic (religious) thinking do not profit from good education as much as non-dogmatic students do. (Schillinger 2006; Lupu 2009; Saeidi 2011; Liaquat 2013; Akin 2018) The question, whether students with dogmatic thinking can profit from special methods like the KMDD, needs to be more systematically researched. My impression from many years of using this method in many different cultures and subcultures is: yes they do.

(c) Hypothesis: Methods of fostering moral competence like the KMDD have *wished-for side-effects, but no unwished-for side-effects*.

Findings: My long-time experience confirms this hypothesis. The KMDD has strong wished-for side-effects like a better learning climate but does not seem to have any unwished-for side-effects. (Lind 2002; 2016)

(d) Hypothesis: Fostering moral competence with methods like the KMDD is very *efficient*. That is, the costs of teaching, implementing and using the method are low as compared to the foreseeable big benefits of this method.

Findings: There are no exact calculation yet which could be used to test this hypothesis. Yet I am very sure that a slight increase of every citizen’s level of moral competence, that would cost little money, would vastly improve the quality of social life and thus safe

society much more money: it would reduce damages due to criminal behavior, and lower the huge costs for controlling, catching and punishing criminals. (Lind 2016)

If time and money are scarce, it is better to focus on a narrow research question, instead of trying to answer numerous questions only superficially. Other researchers can take up the questions which you leave out – if you communicate your study properly.

Science is a collaborative project of many people – lived, living, and to live

Scientific discoveries are rarely the work of a single scientist, but are mostly the result of the cooperation between many scientists: with those who live, those who had lived and those who will live. Scientists do not work only for themselves but for mankind, especially for practitioners who rely on reliable and understandable knowledge. Therefore, good scientists need to *strive* for simplicity of their theories (their complexity develops by itself), and also for clarity, information value, efficacy, and efficiency.

Simplicity: Don't be afraid of simplicity. It is better to focus only on one narrow research question and do this well, instead of attempting to cover "everything." It is better to start with simple theories and hypotheses instead of complex theories. (Complexity is not a virtue, nor is it a trait of the objects which we study. It rather describes our level of ignorance. Anything looks complex as long as we do not understand it.) "Do not multiply concepts without necessity!" This basic principle of good science is attributed to the 14th century philosopher and monk William of Occam. Great natural scientists like Bertrand Russell and Albert Einstein have reinforced it. If they have the same predictive power, simple theories should be preferred over complex ones because they are testable. Without this basic principle we would not have modern natural science or modern technology. Simple theories can be turned more easily into applications. Not to speak of education: Simple theories can be more easily taught to students than "complex" ones. Unfortunately, in social science, the opposite, complexity, is still considered a virtue.

Clarity: Be clear in your thinking and your writing. Clarity is as important for good science as simplicity. In science clarity means essentially that you report in detail the operations of your observation (and measurement) through which you identified your object (e.g., the trait which you studied). This is the reason why a science needs to develop its own language, that is, its own terminology, grammar, and pragmatics.

Unambiguous terminology: In everyday life, and especially in poetry, we often use the same word for different things, and we call the same thing different names. For example, the word "he" can mean every male person in the world. But when I am talking with my wife about a certain neighbor, she clearly understands whom I mean when I say "he." The context made "he" unambiguous. Or, to take another example, the word "wave" means something quite different when we use it to refer to a water wave, an air wave, an electrical wave or a fashion wave. Ambiguity also arises from the fact that a certain object can be named in different ways: a woman is referred to by her husband as wife, by her children as mother, by her sibling as sister, by her grandchildren as grandmother. Mostly we can live with this ambiguity of meaning. Mostly, the context makes it clear what a certain word means, and that the different names refer to the same object—but not always. In some instances the lack of clarity can lead to confusion, or even conflicts and fights. Saying "*he* is stupid," can trigger harsh reactions of someone who is not meant with "he." In psychology we have made little, if any progress, over the past decades, not because they it is a young science but, as the philosopher Ludwig Wittgenstein (1953) has observed "in psychology there are experimental methods and conceptual confusion."

Definitions: Especially in a field like psychology in which no standard terminology is enforced, it is necessary that researchers clearly say how they define their core concepts. They rarely do so. In many cases they do not even tell the readers their definitions at all, nor do they refer to literature in which one can find the definitions. And if they define their concepts they do not tell the reader how their concept is to be observed or measured. Instead they often arbitrarily choose among several ways of observation letting the reader confused about the reasons for choosing this instead of other operational definition. Or they give reasons for their choice which have nothing to do with the validity of the observation. To make things even worth, some psychologists (and other social scientists) use different terminology for the same object. “This is very confusing” students often complain. They are right. The only way to find out how a certain trait is defined is by looking closely at what method of observation is used.

In sum: Always state the definitions of your core concepts clearly. Describe them in detail. Tell the reader through which operations you have observed the trait under study. Use your terminology consistently. Tell the reader if you have changed names or meanings of concepts and why you did so. Avoid using the same name for different objects (as we commonly do in normal life). An avoid using different names for the same object (as poets like to do).

What topic should I choose?

This is the next important question. The answer is up to each researcher. But I can tell you some guiding questions which I use myself:

- (a) Which human trait deserves, or needs, to be researched? (Remember: Try to focus on a single trait or on a very few traits. This is enough work.)
- (b) Is this trait clearly defined, or do you think that it can be scientifically defined in the way I described above? If the answer to both questions is No, you’d better think your plans over. I have looked into many trait concepts (like intelligence, motivation, empathy) which have, at some time, been hotly discussed in psychological journals. But when I could not detect a clear, commonly shared definition or clear description, I decided not to spend time on it any longer.
- (c) What common beliefs and which scientific theories exist about the *relevance* of that trait? Which of these beliefs or theories are so important that they should be submitted to scientific scrutiny?
- (d) What beliefs or theories exist about the means and methods for *fostering* that trait? (Continue like point c).
- (e) Are there two competing hypotheses A and B which you can compare with each other? It is more rewarding when we find out that method A is more effective and efficient than method B, or vice versa, instead of merely rejecting a singular hypothesis. (Of course, more than two hypotheses can be compared, too.)

I have given you here some ideas. But you may have your own ones. In the appendix I present the list of research questions which I have written up for myself.

How should I name the object of my research?

There are many names for the same trait. When you take a closer look at studies of apparently different traits like attitudes, values, orientations, preferences, positions, affect, etc. you find that all are measured in very similar ways through Likert scales. The same is true for traits like competence, ability, capability, and cognition. Their common feature is that they contain certain tasks which are not easily solved. So do not let yourself be confused by different

names – and do not add to this confusion. You should choose a proper name for the trait which you want to focus on, and then define and describe it. Do not introduce a new one if a proper name already exists. For anyone seriously interested in your research topic it is awful if one has to do linguistic studies first in order to find out what you really mean.

Choose the name carefully so that you do not need to change it soon again. This will be the “brand” name by which your research and you will be recognized. In contrast, the description of the trait under study may need improvements over time. The improvement of a description or even of the definition can be, and in fact often is, an important outcome of research. For example, from the beginning I have called the object of my research *moral competence*. (Lind 1978) However, I had also used the Kohlbergian term “moral *judgment* competence” and had called my test “Moral Judgment Test” (MJT). When I discovered that this confused readers, I have dropped the term *judgment* and renamed my test to “Moral Competence Test” (MCT).

Name of the trait: Better than giving you some abstract ideas, I tell you what considerations have led me to choose “moral competence” as name for the trait on which I am focusing on for the past forty years. There were many other options: morality, moral orientation, moral reasoning, moral judgment, virtue, character, civic competence, social competence, emotional competence, values, and many more. Today many of these concepts are more popular than the term “moral.” Yet the term “moral” has the longest tradition. “Moral” issues have been most profoundly researched and discussed since the early Greek philosophers Socrates, Plato and Aristotle. All non-natural sciences – economics, sociology, law, philosophy etc – were once called “moral sciences.” If we would give up this term in favor of more recent terms, we would make it difficult for future generations to connect to this great wealth of literature. (I must confess however that my choice made it very difficult for several years to get anything published.)

I have chosen the particular meaning of “moral” which scholars like Immanuel Kant and Lawrence Kohlberg attached to this term: “to act upon moral principles.” Others use this term frequently to mean “to act upon the orders of some authority.” However, such behavior is better described by terms like “obedience,” “conformity,” or “law-abiding.” It would be very confusing to use the word “moral” to mean both kinds of behaviors.

As I have shown above “moral behavior” must be described by two distinct aspects: moral orientations and competence. (Lind 2016)

Definition of the trait: The definition of a trait must allow anyone to identify the trait clearly and unambiguously. It should consist of terms and operations which are familiar to the readers. They should not consist of terms which need to be defined themselves. To be clear and unambiguous, the definition should also state what the trait does NOT mean. For example, I define moral competence as *the ability to solve problems and conflicts morally, that is, on the basis of general moral principles through deliberation and discussion*. In order to make participants’ moral competence visible in a test, they are asked *to rate arguments pro and contra a certain solution of a conflict or problem, whereas the arguments represent six different types of moral principles*. The moral competence of people manifests itself in the degree to which their response pattern is determined by *the moral quality of the arguments instead of the arguments’ opinion agreement*.

Moral competence does not mean the ability to discuss ethical concepts or apply formal logic to ethical statements. This is better called ethical competence. Neither does moral competence mean the ability to obey commands of others. And it does not mean some “latent trait” which can be inferred from test responses of many people.

Description of the trait (its nature): When we speak of the “nature” of a trait we do not mean that this trait is inborn or unchangeable, as is often said. Rather we mean by nature the sum of

what we know about it. As we have seen, we already know a lot about the nature of moral orientations and moral competence.

Testing theories with experiments

Science is in some respect like a sewage plant. The unique job of scientists is reduce information pollution, namely to separate false theories from correct ones through rigorous experiments. Ridding mankind from a false theory is as valuable as to find a new theory. Scientists are most happy, however, if they can not only refute a false theory but also replace it by a better one.

Our two most powerful instruments for singling out bad from good theories are (a) analyzing the information value of their hypotheses, and (b) doing rigorous experiments. Profound treatments of these two concepts—theories (or hypotheses) and experiment—fill big volumes. I think that they can be distilled to some simple truth which I think will improve psychology as a science:

Hypotheses have a probability value

Formulate hypotheses in such a way that they have a low probability of being confirmed merely by chance. Low probability of confirmation, also called “refutability,” of hypotheses is considered to be the most important criterion for demarcating science from non-science. Note that *probability* is a term of logic, not of empirical frequency. The lower a hypothesis’ probability, the higher is its information value, that is, the more interesting is it to invest time and money into an experimental test of their truth. (Popper 1968)

This feature of good science rules out hypotheses like “moral competence *differs significantly* between males and females.” This prediction has no information value because it is always true ($p = 1.0$). Avoid also predictions of this kind: “The moral competence in group A is larger than in group B,” because the chances are still very high (50% or $p = 0.5$) that it will be confirmed by data. The chance of confirmation increases to 100% if you predict “that moral competence is bigger in group A than in group B in a series of experiments.” This is as if would you predict that in a series of dice rolling the number 1 will show up in $1/6^{\text{th}}$ of the cases. This is always true. You do not need to study this hypothesis through rolling dices ten thousand times. The hypothesis “Variable A *may* correlate with variable B” has also zero information value. It does not say how high the correlation should be and it does not exclude a non-correlation. Hence, it predicts that everything is possible ($p = 1.0$).

Yes, it takes a little courage to formulate informative hypotheses with a very small probability. But science flourishes only if scientists have that courage. A good example for a courageous hypothesis with a high information value is Kohlberg’s Stage theory. Based on this theory he predicted that people’s moral judgment competence proceeds from Stage 1 to Stage 6 on his scale of moral development. The probability that a particular participant’s development confirms this hypothesis is much smaller than the conventional significance-level of $p=0.05$, namely $p=0.0014$ because for six Stages there are $6 \times 5 \times 4 \times 3 \times 2 \times 1$ (in short: factorial of six $6!$) = 720 developmental sequences possible. Therefore, the chance of a particular sequence is $p=1/720=0.0014$. Since Kohlberg’s theory lets us predict this sequence for many people, say for 100 people, the probability of confirming our model by chance in a study of 100 people, becomes extremely small, namely $0,0014^{100}$. What a courageous hypothesis! This kind of scientific has spurred experimental research. It has finally let us to refute Kohlberg theory of invariant Stages, and to replace it by our Dual-Aspect-Theory: moral orientations are invariant, but moral competence can regress. (Lind 2016)

We have used this logic also for predicting that Kohlberg's six types of moral orientations form a certain order of preference (from type 1 to type 6). For a single person this probability is also $p = 0.0014$. If we study 100 persons' moral orientations, the probability of confirming our model by chance is extremely small, namely $0,0014^{100}$, which is far beyond the conventional significance level $p=0,05$. In spite of this extremely low probability, this hypothesis is, as we have seen, well supported by experimental data. Only small deviations are found, namely small inversions of preference between Type 1 and 2 orientations, and between Type 5 and 6 orientations. (Lind 2016)

In sum, theories should be as precise as possible, whereby precision means much more than many digits behind the decimal point. Such pseudo-preciseness is not a proof of good science. More important is a precise use of words and terminology.

Theories about the nature, relevance and changeability of psychological traits should be of psychological, not statistical nature. *Obviously, good psychological theories lets us formulate hypothesis with a much higher information value than the widely used statistical $p=0.05$ hypothesis. Of course, statistical models are useful for psychological applications in education, therapy, and leadership only if they are based in psychological theories.*

Definition of Experiment

An experiment is any carefully arranged situation in which a hypothesis is put to rigorous test so that the outcomes let us unambiguously confirm or disconfirm that hypothesis. (Popper 1968)

Contrary to wide-spread belief in psychology and the social sciences, an experiment does not need to contain manipulation of a variable. One of the most famous experiments in physics, in which Einstein's relativity theory was put to test, was arranged as an observation. The physicist Eddington tested Einstein's hypothesis that the mass of the moon would diffract the rays of the sun only by observation. (Dyson et al. 1920) He could not move the moon in front of the sun but could wait only until the moon moved by itself before the sun.

Nevertheless, the manipulation of a supposed factor is a very powerful method to test a causal hypothesis.

In psychology the use of experiments is more limited than in the natural sciences because we have to respect the dignity of our research "objects," and privacy protection laws.

Nevertheless, we should do more experiments. But note: good experiments often require much time for preparation and a good bit of fantasy and ingenuity.

Testing theories with correlation studies

Conventionally, correlation study means that relevant data are collected only through observation or measurement only, but not through the manipulation of one or more factors. Yet, as we have seen, the decisive difference between an experiment and a correlation study is the quality of the research design. There are very good experiments in which no variable has been manipulated.

The aim of a good experiment is to produce data that let us either *unambiguously confirm* or *unambiguously refute* the hypothesis in question. Therefore, an experiment must be very *carefully* planned and executed. Regardless of whether or not some factors need to be mani-

pulated, an experiment usually costs much more time, money and brain than a correlation study.

Correlation studies are done when there is less time, money, or brain available than would be needed to test a hypothesis. Because correlation studies invest less into the design and performance of a study, they usually produce ambiguous data which do not clearly support or refute the hypothesis in question. To compensate for this drawback, data of correlation studies are submitted to more or less complex statistical analyses. Their aim is to reduce ambiguity and help the researcher to decide on the validity of their hypotheses.

However, statistical methods like correlation statistics and its derivatives as substitutes for good experimental design hardly make up for bad research design.

The most commonly used correlation statistics, the *Pearson coefficient of correlation* r_{xy} varies from -1 to +1. If a study finds that the correlation r between gender and hair color is zero, this means that these two variables are not statistically correlated. In other words, there are as many blond men as there are blond women. $r = 1.0$ means that both variables are perfectly correlated. This could mean that *all* men are blond and all women are dark-haired (or vice versa). Does statistic unambiguously prove any causal relationship between gender and hair color?

The answer is no. Statistics by itself cannot prove or disprove a causal relationship because statistics only count phenomena but cannot reveal causal relationships. There can be spurious correlations and spurious non-correlations. And there can be non-linear correlations which produce a zero r .

The web-site <http://www.tylervigen.com/spurious-correlations> gives 30.000 funny examples for spurious correlations based on real data. For example, it shows that the per capita consumption of mozzarella cheese is highly correlated with the number of civil engineering doctorates awarded on the US. So eating cheese will help you getting a doctorate.

This, of course, is funny nonsense. Yet it is not so funny when researchers actually base their theories and their recommendations for health care, education, good policy-making etc. on such spurious correlations and the media uncritically publicizes them. Usually such spurious correlation coefficients are produced by some third variables which directly or indirectly determine the two variables which correlate.

There also spurious non-correlations, that is, zero correlation coefficients when actually there is a causal relationship. A famous example is the Yerkes-Dodson (1908) law which predicts that up to a certain point habit-formation (performance) goes up with stimulus-strength (excitement), but beyond that point the relationship becomes negative: performance goes down again with growing excitement. This law is well supported by many experiments. But when one calculated a coefficient of correlation, the outcome is zero.

Nevertheless, the outcomes of experiments are sometimes also analyzed statistically with correlation statistics or similar methods. But this does not make them correlation studies. Again, it is the design of the study which makes the difference, not the use or not-use of correlation statistics.

In many statistics textbooks and research studies, the calculation of coefficients of correlation (like Pearson's r) is mystified. They say that it must only be calculated if the variables (e.g., gender, intelligence, moral competence) are distributed "normally." This is pure nonsense. It

is true that the size of r is affected by the shape of the distribution of the variables, but it is also affected by many other factors which cannot be easily teased out. But this does not make this statistics inapplicable. Even worse, many believe that this presumed handicap can be cured by “normalizing” the distribution of human traits mathematically, which means that they squeeze and stretch any distribution until it is “normal.” This exercise has not scientific justification.

Listen to Karl Pearson who invented the term “normal” to describe the “distribution of errors” (not of human traits!). Pearson writes: “Many years ago I called the Laplace-Gaussian curve the normal curve, which name, while it avoids an international question of priority, has the disadvantage of leading people to believe that all other distributions of frequency are in one sense or another 'abnormal.' That belief is, of course, not justifiable. It has led many writers to try and force all frequency by aid of one or another process of distortion into a 'normal' curve.” (Pearson 1920, p. 25) With “international question or priority” Pearson alludes to the fight over the question who truly invented this curve, the French Laplace or the German Gauss. So this unfortunate name for the distribution of errors misled generations of researchers to falsely believe that every human trait was “normally” distributed, and if not, had to be forced into an error distribution.

In sum, it is worthwhile to invest time, money and brain into a careful experimental design, instead of hoping for the wonders of statistical correlations or applying statistical voodoo magic.

Good research reporting

You should give sufficient information so that the reader can understand your research question, hypotheses and findings – in psychological terminology. Report statistical models, methods and findings only as much as is necessary. In too many research reports too much statistical things are reported, and too little psychological. Moreover, in too many reports, the statistics are overly complex and little understood.

Do not forget to report mean scores and medians, and display them graphically. Transformation of data makes them unintelligible and incomparable with other studies. Mostly these transformations are not justified or needed, but are reported only because editors and reviewers request it for conventional reasons.

Report absolute effect sizes. Statistical significance is no substitute for this. A simple measure for effect size is the difference between the means of posttest and pretest scores in an experimental group “E”:

$$E_2 - E_1 = aES.$$

If a comparison group “C” exists, the absolute effect size is defined as the difference of two differences:

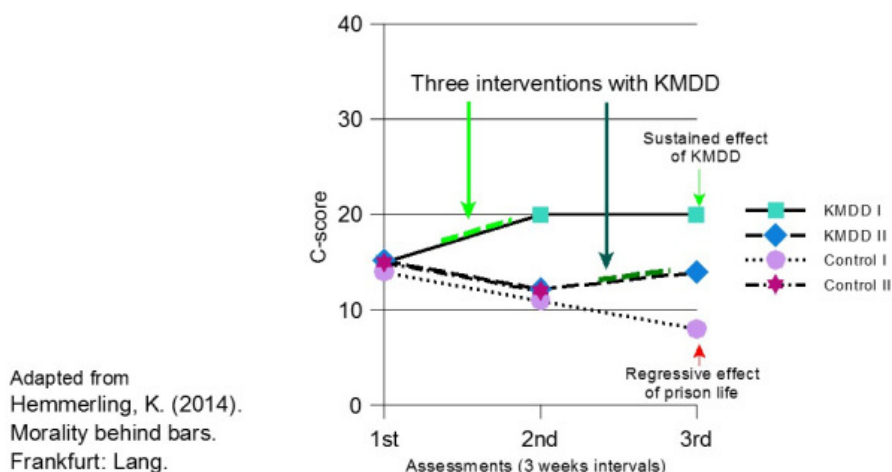
$$(E_2 - E_1) - (C_2 - C_1) = aES.$$

In contrast, tests of statistical significance either vastly overestimate the effects of a treatment (“highly significant beyond the 0.001 level”), or overlook the effects of treatments (“missed the $p=0.05$ level”), depending on the size of the sample.

Coefficients of correlation “ r ” and Cohen’s “ d ” are measures of *relative* effect size. They are relative because they depend on the variance of data in a particular sample. Hence they cannot be compared across samples or studies with different variances (see also above).

Graphics should be self-explaining: Do not use abbreviations (except very common ones). Denote the axis and print the numbers which you graphically depict, so that the readers do not have to scan your whole text for this information. Do not spread out your y-axis in order to make a small difference look big. For C-scores the axis is conventionally set to 0 to 40,

unless higher scores are found. See graph of the intervention experiment with prisoners by Hemmerling (2014).



Tables should also be self-explaining and not too large. If large tables with many numbers are necessary they should be placed in the appendix.

Use quotes to honor a scholar who made an important contribution to the field of your research. Do not use quotes in place of convincing arguments and data (“borrowed authority”).

Use references to inform readers where they can find background information. Do not use references in place of a definition and description of your research topic.

If information is central for your study then you should use a whole sentence or even a whole paragraph to describe it. Your readers will be grateful. Do not put two and more important information into one sentence or into a single adjective. (I have seen this being done often!)

Virtues of research practice

- Be focused. Your study should focus on testing only *one main* hypothesis or prediction. You may also examine some auxiliary hypotheses which are closely related to your main hypothesis. But not too many! Your readers and editors will also appreciate a clear focus.
- Be in charge. You must take responsibility for your research question and research findings. This means that you should take authorship for your main predictions and your examination of their validity. This means also that you should be able to make recommendations and give examples for your recommendations.
- Be fair. Do not attack theories which you do not really know. Do not paraphrase their core concepts and statements in a loose or even sloppy manner. Quote always the most recent publications about these theories.
- Do not ignore important research literature just because it has been published many years ago. It can be much better than recent research.
- Translation. Remember that words often have different meanings for lay-people and scientist, and that their meaning also varies across disciplines. Especially the term “moral” has many meanings in every-day usage and science. Use theoretical terms cautiously in everyday contexts or else your audience will misunderstand you.

- Use follow-up studies. If you make recommendations about lasting effects you should be able to back them up by findings from follow-up studies. Follow-up studies should be done at least a couple weeks after the post-test, but should not be done too much later because interventions can hardly be expected to have “eternal” effects. Most effects can fade out.
- Protect privacy. If you do pretest-posttest comparison and follow-up then you must find a way to protect the privacy of your participants. Don’t forget to give them a good explanation why you want them to fill out the same test(s) twice or more often, otherwise they may not give you a sincere answer or refuse to respond at all. Protect also the privacy of the participants’ institution and its staff. If you want to reveal the participants’ name or institution you must get their permission.
- Compare findings. You can appreciate the effects of interventions only if you compare the changes in the experimental group with changes in other groups. For example, there can be a strong effect even though there was no or little change in an experimental group. For example, in his study of juvenile prisoners, one of my master students found no change in his intervention group, but at the same time a regression of moral competence in his control group. Hence his intervention had an effect as it stopped regression. Own control-groups are not the only way to compare. You can also compare your findings with the findings of other research (see: <http://www.uni-konstanz.de/ag-moral/b-publik.htm>).
- Randomization overrated. It does not rule out false interpretations. Full randomization is hardly ever possible because many important variables cannot be randomized. You cannot randomize over time. Often one cannot randomize all participants in a study or all; voluntary participants cannot be replaced by mandated participants. One cannot randomize control groups if they are too expensive. Moreover, studies with randomized samples have not shown to be superior to studies with matched samples (Lipsey & Wilson 1993). In sum: There is no mechanical method which can replace the scientist’s own critical judgment and his/her ingenuity in designing experiments that reduce or rule out false interpretations.
- You do not need to re-invent the wheel over and over again. The research field of moral competence has produced many findings that are well established. We have really made some scientific progress! Of course, there are still a lot of questions which are not yet answered or are still debated. But we do not need to start always from the scratch.
- Style of writing / design of report. The first and the last chapter (introduction, conclusion) of your research report (i.e., book) should be written in clear, simple, non-technical language, speaking to a general audience (that is, all people belong to this group besides a very few experts). If you think you should make the reader acquainted with some core technical concepts you must explain them with familiar examples (not with other unfamiliar concepts).
- Last but not least. To say something simple is not simple at all. It took me more than 40 years of research to come up with these simple ideas.

Dedication

I dedicate this paper to my teacher, colleague, and good friend *Roland Wakenhut*, professor of social psychology at the University of Eichstätt, who has passed away much too early. He has been a continuous source of motivation for me. I miss him.

Appendix: Important definitions

Moral judgment competence (Kohlberg 1964)

“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)

Deutsch: Moralische Urteilsfähigkeit ist „das Vermögen, Entscheidungen und Urteile zu treffen, die moralisch sind, das heißt, auf inneren Prinzipien beruhen und in Übereinstimmung mit diesen Urteilen zu handeln.“ (meine Übersetzung)

Source: 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.

The clarity and information value of this definition attracted me to Kohlberg’s work. Herein Kohlberg defines moral competence as an ability (not as a value), as an internal trait (not as an external attribute), and as observable in manifest behavior (not as an unobservable statistical construct).

Five decades of research showed me how fruitful his definition was but also that it needed revision:

Moral competence (Lind 2016, supplemented)

I have defined moral competence as *the ability to cope with problems and conflicts on the basis of moral principles through deliberation and discussion, instead of through violence, deceit, submitting to others, or ignoring.*

Source: Lind, G. (2016). How to teach morality. Promoting deliberation and discussion. Reducing violence and deceit. Berlin: Logos-Verlag.

Appendix:

Unsystematic collection of research ideas (add your own ideas!)

1. Is moral competence relevant for living together in a democracy?
 - a. Does moral competence reduce the use of *violence* when coping with problems, conflicts, dilemmas, etc. in various contexts (e.g., school, ethnic relations)? There seem to be not many studies; see Seitz (1991; www.uni-konstanz.de/ag-moral/mut/mjt-references.htm)
 - b. Does moral competence lower the occurrence of deceit? See the classic study by Hartshorne & May (1928). *Studies in the nature of character*. Vol. I: *Studies in deceit*, Book one. New York: Macmillan.
 - c. Does moral competence keep up the level of cooperation in so-called prisoner dilemma games when the situation is hindering cooperation? (for the many forms of pd games, and the huge literature in experimental economics and psychology, see: <http://www.econlib.org/library/Enc/PrisonersDilemma.html>).
 - d. Does moral competence reduce *pluralistic ignorance* in regard to various phenomena (e.g., helping behavior, use of violence, deceiving, cooperation, competition). See: http://en.wikipedia.org/wiki/Pluralistic_ignorance, and Lind, G. (1995). What do educators know about their students? A study of pluralistic ignorance. In R. Olechowski & G. Svik, eds., *Experimental Research on Teaching and Learning*. Frankfurt, Bern: Peter Lang, pp. 221-243.
 - e. Does moral competence reduce the tendency to submit to others? Does moral competence reduce conformity to authority or to masses? (similar to the experiments by Solomon Ash.)
 - f. Does moral competence facilitate helping behavior? McNamee, S. (1977). Moral behavior, moral development and motivation. *Journal of Moral Education* 7(1), 27-31.
 - g. Does high moral competence stimulate learning of various subject matters? See e.g., Heidbrink, H. (2010). Moral judgment competence and political learning. In: Lind, G., Hartmann, H.A. & Wakenhut, R., eds., *Moral judgment and social education*, pp. 259 - 271. New Brunswick, NJ: Transaction Publisher.
 - h. Does high moral competence help to make quicker decisions? See: Mansbart, F.-J. (2001). *Motivationale Einflüsse der moralischen Urteilsfähigkeit auf die Bildung von Vorsätzen*. (Motivational influence of moral judgment competence on the formation of intentions.) Unpublished master thesis. Dept of Psychology, University of Konstanz, Germany. www.uni-konstanz.de/ag-moral/mut/mjt-references.htm.
 - i. Does moral competence help to prevent drug consumption? See: Lenz, B. (2006). *Moralische Urteilsfähigkeit als eine Determinante für Drogenkonsum bei Jugendlichen*. [Moral judgment competence as a factor of substance consumption in adolescence.] Master thesis, Department of Psychology, University of Konstanz. Download: www.uni-konstanz.de/ag-moral/mut/mjt-references.htm.
 - j. Does moral competence help to engage in, and maintain, a controversial discourse with opponents and enemies?
 - k. Does moral competence help to develop reason in a context which is hostile to reason?
 - l. Your research question: ???
2. How much do school and university promote moral competence in various countries?
 - a. What is the increase of moral competence in one school year in middle school, high school, college, university, vocational school?

- b. Do comprehensive schools, in which there is no segregation by ethnicity and economic status, show more progress of moral competence?
 - c. Do schools with open student-oriented teaching promote moral competence more than traditional schools with teacher-directed learning?
 - d. How can one increase opportunities of responsibility-taking and guided reflection as a teacher and as an educational policy-maker? (See Lind 2002; Schillinger 2006; Lupu 2009; Saeidi 2012)
 - e. Your research question: ???
3. Can moral competence be taught? And how?
- f. Do certain changes of the KMDD improve its efficacy?
 - g. Do certain changes of training KMDD-Teachers improve its efficacy?
 - h. Are there other forms of teaching (which contain basic KMDD principles) that produce the same or even stronger effects?
 - i. Which other effects does the KMDD have? Does it also foster learning motivation?
 - j. Do KMDD-sessions improve the learning climate in the classroom, or the rehabilitation of prisoners, or other attributes of the institution?
 - k. How many KMDD-sessions are needed to achieve a certain level of moral competence in all participants within a certain time range?
 - l. Do KMDD-sessions serve ALL participants in a class, or do they create inequality among the participants (which would contradict its aims)?
 - m. How can regression of moral competence be prevented in various institutions (e.g., medical students, prisoners)? See Lind, G. (2000). Moral regression in medical students and their learning environment. *Revista Brasileira de Educacao Médica* 24 (3), 24-33. And other studies: www.uni-konstanz.de/ag-moral/mut/mjt-references.htm.
 - n. How efficient is the KMDD? What are its costs and do they compare these costs with the savings it produces?
 - o. Your research question: ???
4. How does *segmentation* of moral competence develop and what effects does it have for behavior and social life? Segmentation is defined as a difference of more than eight C-points between the two stories in the *Moral Competence Test*. So far, only in the “doctor dilemma” moral segmentation has been found. Segmentation seems to be connected with confessional religiosity but not with personal religiosity. (Saeidi 2011; Liaquat 2013; Senger 2010; Bataglia & Schillinger, 2013)

References

- Akin, A. (2018). *Religiöse Erziehung und Moralentwicklung. Der Einfluss einer religiös geprägten Kindheit auf die moralische Orientierung und Kompetenz*. Bachelorarbeit in Psychologie. Hochschule für Gesundheit und Medizin, Berlin.
- Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.
- Althof, W., Garz, D., & Oser, F., eds. (1999). *Moralisches Urteil und Handeln*. Frankfurt: Suhrkamp.
- Bataglia, P. & Schillinger, M. (2013). Moral segmentation in studies with the Moral Judgment Test in Brazil. In: E. Nowak, D. Schrader & B. Zizek., eds., *Educating competencies for democracy*, pp. 71-82. Frankfurt am Main: Peter Lang Verlag.
- Bortz, J. (1983). *Lehrbuch der empirischen Forschung für Sozialwissenschaftler*. Berlin: Springer-Verlag.
- Carver, R. P. (1993). The case against statistical significance testing, revisited. *Journal of Experimental Education*, 61(4), 287-292.

- Cohen, J. (1994). The Earth Is Round ($p < .05$). *American Psychologist* 49 (12), 997-1003.
- Colby, A. & Kohlberg, L. (1987). *The measurement of moral judgment*. Volume I and II. New York: Columbia University Press.
- Dyson, F. W., Eddington, A. S., Davidson, C. (1920). A Determination of the deflection of light by the sun's gravitational field, from observations made at the total eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London*. Volume 220, pp. 291-33.
- Emler, N., Renwick, S. & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology* 45, 1073-80.
- Hemmerling, K. (2014). *Morality behind bars – An intervention study on fostering moral competence of prisoners as a new approach to social rehabilitation*. Frankfurt: Peter Lang.
- Jacobs, M. K. (1975). *Women's moral reasoning and behavior in a contractual form of prisoner's dilemma*. Toledo, OH, University of Toledo. Unpublished doctoral dissertation.
- Liaquat, A. W. (2013). *Effect of dogmatic religiosity and educational environment on moral judgment competence*. Unpublished dissertation, Dept of psychology, International Islamic University Islamabad.
- Lind, G. (1978). Wie misst man moralisches Urteil? Probleme und alternative Möglichkeiten der Messung eines komplexen Konstrukts. [How does one measure moral judgment? Problems and alternative possibilities of measuring a complex construct]. In G. Portele, ed., *Sozialisation und Moral*, pp. 171-201. Weinheim: Beltz.
- Lind, G. (1985). *Inhalt und Struktur des moralischen Urteils. Theoretische, methodologische und empirische Untersuchungen zur Urteils- und Demokratiekompetenz bei Studierenden*. [Content and structure of moral judgment. Theoretical, methodological and empirical studies on judgment and democratic competence of university students.]. Doctoral thesis. Social Science Faculty, University of Konstanz.
- Lind, G. (1986). Parallelität von Affekt und Kognition in der moralischen Entwicklung. [Parallelism of affect and cognition in moral development]. In F. Oser, W. Althof & D. Garz, eds., *Moralische Zugänge zum Menschen - Zugänge zum moralischen Menschen*, pp. 158-179. München: Peter Kindt.
- Lind, G. (1989). Essay review: 'The measurement of moral judgment' by Anne Colby, Lawrence Kohlberg. *Human Development* 32, 388-397. [PDF](#)
- Lind, G. (2000). The importance of responsibility-taking opportunities for self-sustaining moral development. *Journal of Research in Education* 10(1), 9–15. [PDF](#)
- Lind, G. (2002). *Ist Moral lehrbar?* Berlin: Logos.
- Lind, G. (2008a). 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. [PDF](#)
- Lind, G. (2008b). 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. [PDF](#)
- Lind, G., Sandberger, J.-U. & Bargel, T. (2010). Moral competence and democratic personality. In G. Lind, H.A. Hartmann & R. Wakenhut, eds., *Moral Judgment and Social Education*, pp. 55-77. Rutgers, NJ: Transaction Books, second edition. (First edition 1985 by Precedent Publishing).
- Lind, G. (2011). Editorial: Moral competence and the democratic way of living. *Europe's Journal of Psychology* 7, 5, 569 - 596.
- Lind, G. (2016). Effektstärken: Statistische versus praktische und theoretische Bedeutsamkeit. Erweiterter Vortrag. [Effect sizes: statistical, practical, and theoretical significance of empirical studies.] [PDF](#)
- Lind, G. (2016). *How to teach morality. Promoting deliberation and discussion. Reducing violence and deceit*. Berlin: Logos-Verlag.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist* 48, 1181-209.

- Lupu, I. (2013). Moral competence and dogmatic religiosity. In: E. Nowak, D. Schrader & B. Zizek., eds., *Educating competencies for democracy*, pp. 379 - 390. Frankfurt am Main: Peter Lang Verlag.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin* 105(1), 156-166.
- Mofakhami, A. (2018). The effect of moral competence on conformity behavior. Master thesis. Allameh Tabatabaei University Teheran.
- Nowak, E., Schrader, D. & Zizek, B. eds. (2013). *Educating competencies for democracy*. New York: Peter Lang Verlag ([table of contents](#)).
- Pearson, K. (1920). Notes on the history of correlation. *Biometrika* 13, 1, 25-45.
- Piaget, Jean (1976). The affective unconscious and the cognitive unconscious. In: B. Inhelder & H. H. Chipman, eds., *Piaget and his school*, pp. 63-71. New York: Springer.
- Popper, K. R. (1968). *The Logic of scientific discovery*. London: Hutchinson (Original 1934).
- 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. (1978). *Development in judging moral issues*. Minneapolis, MI: University of Minnesota Press.
- Saeidi-Parvaneh, S. (2011). Moral, Bildung und Religion im Iran – Zur Bedeutung universitärer Bildung für die Entwicklung moralischer Urteils- und Diskursfähigkeit in einem religiös geprägten Land. [Morality, education and religion in Iran. On the importance of university education for the development of moral judgment and discourse competence in a religiously shaped country.] Doctoral dissertation, Dept. of Psychology, University of Konstanz.
- Schillinger, M. (2006). *Learning environments and moral development: How university education fosters moral judgment competence in Brazil and two German-speaking countries*. Aachen: Shaker-Verlag. [PDF](#).
- Senger, S. (2010). Segmentation of soldiers' moral judgment. In G. Lind, H. A. Hartmann & R. Wakenhut, eds., *Moral judgments and social education*, pp. 221-242. New Brunswick, NJ: Transaction Publisher.
- Shaver, J. P. (1993). What statistical significance testing is, and what it is not. *Journal of Experimental Education* 61, 293-316.
- Wittgenstein, L. (1953). *Philosophical investigations*. Translated by G.E.M. Anscombe. New York: The Macmillan Company.
- Yerkes R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*. 18, 459-482

Source of article:

http://vg01.met.vgwort.de/na/58ccb7c4138e472a955e3b5c587a71bc?l=https://uni-konstanz.de/ag-moral/pdf/Lind-2019_The_Art_of_Experimental_Moral_Psychology.pdf

For a more extensive list of literature visit:

Moral psychology: <http://www.uni-konstanz.de/ag-moral/mut/mjt-references.htm>
 Moral education: <http://www.uni-konstanz.de/ag-moral/moral/kmdd-references.htm>
 Complete list of my publications: <http://www.uni-konstanz.de/ag-moral/b-liste.htm>

© Copyright by Georg Lind