Testing the moral algebra of two Kohlbergian informers

Wilfried Hommers*, Martin Lewand, Dominic Ehrmann

University of Würzburg, Germany

This paper seeks to unify two major theories of moral judgment: Kohlberg´s stage theory and Anderson´s moral information integration theory. Subjects were told about thoughts of actors in Kohlberg's classic altruistic Heinz dilemma and in a new egoistical dilemma. These actors's thoughts represented Kohlberg's stages I (Personal Risk) and IV (Societal Risk) and had three levels, High, Medium, and Low. They were presented singly and in a 3 x 3 integration design. Subjects judged how many months of prison the actor deserved. The data supported the averaging model of moral integration theory, whereas Kohlberg's theory has no way to handle the integration problem. Following this, subjects ranked statements related to Kohlberg's first four stages in a procedure similar to that of Rest (1975). Higher score went with larger effect of Societal Risk as predicted by Kohlberg's theory. But contrary to Kohlberg's theory, no age trends were found. Also strongly contrary to Kohlberg's theory, effects of Personal Risk (Stage I) and Societal Risk (Stage IV) correlated positively.

KOHLBERGIAN MORAL ALGEBRA

This paper reports about further steps towards unifying two major theories of moral judgment: Kohlberg´s stage theory (Kohlberg, 1976) and the theory of information integration (Anderson, 2008). Its main purposes are testing the averaging model of the Information Integration Theory (IIT) regarding moral judgment about Kohlbergian concepts as well as testing the influence of a traditional non-interview measure for the Kohlbergian concepts.

Kohlberg´s stage theory is represented by the moral categories obtained with Kohlberg´s Moral Judgment Interview (MJI) while avoiding limitations of Kohlberg´s interview approach. One limitation is that it has been extensively criticized on its demands on verbal facility (Rest et al.,

1999, chapter 2). As Rest (1986, p. 462) noted "production measures require verbal expressiveness in order for the subject to be credited with a cognitive structure." Another limitation of the interview approach is that it misses to study directly the integration of the diverging multiple determinants contained in the dilemmas and interview protocols. And although the stage theories recognize the importance of integrating multiple determinants of moral judgment, they have no way to analyze the integration process.

The theory of information integration offers a frame for studying the integration problem. As previous empirical work has found (e.g. Anderson, 1996) moral judgment frequently depends on the integration of more than one piece of information which followed frequently simple algebraic rules. Even children are sensitive to differing levels of a given moral informer and are capable of integrating them (Anderson & Butzin, 1978; Leon, 1982, 1984).

In similar manner a novel task is employed to unify both theories. The novel task makes low demands on verbal facility. Unlike in the interview approach, the novel thought task presents moral informers as thoughts of the acting agent. The thoughts present typical content of different Kohlberg´s stages. Furthermore, the thoughts are presented in a stimulus design. The moral informers of different content differ in value, for example high and low, and may be combined. The critical prediction of Kohlberg's theory is that differences in effect of those pieces of information should be substantial for a person at a specific moral stage, but small for a person at a different stage.

In this study two contents are employed. Thoughts about the Personal Risk of being caught represent Kohlberg´s first preconventional stage, stage I, heteronomous morality. Thoughts about the Societal Risk, when everybody would act similarly, represent the second conventional stage, stage IV, social system and conscience, as characterized by Colby, Kohlberg et al. (1987, p. 18). Those moral informers are varied and combined systematically in order to study the integration of multiple determinants in moral judgment when the subjects rate how much punishment the actor deserves after they heard about his thoughts when acting.

The second purpose of the present study is concerned with the relation of the IIT approach to another non-interview measure of Kohlbergian morality which may be helpful for advancing the unification of both theories. Therefore, a German non-interview measure, the Würzburg Moral Inventory (WMI), was developed (Lewand & Hommers, 2001). Like Rest´s Defining Issues Test DIT (e.g. Rest, 1975) the recognition task of the WMI only asks for rankings of prototypic moral content of the stages. The

specific question of this part of the study was how individual differences in the WMI affect the effects of the stimuli in the information integration design.

Of course generalisation is of concern in research on morality. Results should be independent of the specific moral problem at least to some extent. Therefore, two cover stories are employed to test the generality of the results. One is the well-known dilemma about Heinz and his sick wife, the other is about Paul who loves expensive watches but lost his purse (Hommers & Lewand, 2005). Both, Heinz and Paul, commit burglary to a store, but from different motives. Heinz acts from altruistic motives of doing something for his sick wife, whereas Paul´s motive appears egoistical. Of course one would expect that the levels of punishment would differ, but the main question is whether the judgmental structures would do so when examined with the IIT approach of the novel task.

Finally, the design should provide the necessary fit to other aspects of the Kohlbergian research. Therefore, subjects of a large age range (9 years to 27 years) are employed, since the development of morality across this age range is Kohlberg´s topic.

METHOD

Two experiments were run as a "judge play", one about Heinz and one about Paul. The tasks were identical except the background scenario, i.e. the cover stories about Heinz and Paul.

Dilemma. At first Kohlberg´s well-known moral dilemma about Heinz was presented: Heinz broke into a drug store and stole the drug after the druggist refused the drug to Heinz because Heinz could not pay for it.

The new vignette of the egoistically motivated Paul told that Paul loves expensive watches but lost his purse with his saved money just before buying the most loved watch. As a consequence he decided to steal the watch. Paul as well as Heinz were caught due to the security equipment of the drug store.

Thought scenario. The thoughts of Heinz and Paul when stealing formed the stimuli of the information integration tasks which were given after a training phase as is the common methodology for testing the averaging model of the information integration theory.

The content of the thoughts were from Kohlberg´s stage I and stage

IV. The stage-I stimulus variable, Personal Risk, had the following three conditions: "The risk of being caught and severely punished is low/medium/high". The stage-IV stimulus variable, Societal Risk, had the following three conditions: "If everybody acted like me, law and order would be at low/medium/high risk in the long run". Those thoughts were presented singly and in combination following the Heinz or Paul cover story. Thus, the combined stimuli were from a 3 x 3 design, e.g. "Heinz thought: The risk of being caught and severely punished is low (or medium/high respectively). If everybody acted like me, law and order would be at high (or low/medium respectively) risk in the long run."

Non-interview measure of morality. The Würzburg Moral Inventory (WMI) was presented after the Heinz and Paul task (Lewand & Hommers, 2001). The four dilemmata of Form A of Kohlberg`s Moral Judgment Interview were used in the WMI. For each choice among the Pro or Contra action five prototypic arguments representing the first five stage levels were given originally. These were formulated in German and with sufficiently distinct content referring to the definitions and examples of the MJI Manual (Colby, Kohlberg, et al., 1987). For the present study the WMI was reduced to its prototypic statements of the first four stages.

Subjects indicated the rank order of the four arguments for each of the eight possible pro-contra choices for the four dilemmata. The ranks were transformed, so that lower summed ranks result in a higher WMI-score representing a higher Kohlbergian stage level.

Reliability analysis of the sum of those transformed ranks by Cronbach´s Alpha with a larger sample (N=420) showed a maximum of reliability (α=.65) for the combined ranks of Stage III and Stage IV arguments (Ehrmann, 2010). Those WMI-scores could principally vary between 24 and 56 and correlated with age, r=0.49.

Procedure. The subjects were to take the perspective of a judge and were informed about their rating task in three steps similar to standard integration-theoretical manner (Anderson, 2008). First, they were introduced to the 13 levels graphical rating scale by giving an initial judgment on the dilemma without added thoughts. They were told to assume a criminal code with applicable imprisonment ranging from 0 to 12 months as the possible judgments.

Second, after this initial judgment thoughts of the actor were given in a list. The subjects were to choose those thoughts which they considered morally demanded of Heinz or Paul when acting. The list included Personal Risk and Societal Risk among other morally relevant or irrelevant content

like window exposition, what friends would do, offering recompense, assuming damages paid by insurance.

Third, punishment ratings of single levels of thought informers were obtained as training. In addition to each of the three levels of Personal Risk and Societal Risk two further thoughts were presented in both experiments: (a) a non- moral thought indicating whether Heinz or Paul liked "the window´s exposition of the store"; (b) an indication about whether the friends of Heinz or Paul would do the same (Kohlberg, 1976; Colby, Kohlberg, et al., 1983).

Finally, after those three steps of the instruction the 3x3 combinations of Personal Risk and Societal Risk were given intermixed with the three levels of each risk variable presented singly to allow for testing the integration rule.

After completion of the Heinz experiment the Paul experiment followed in the same manner. Finally the non-interview measure of morality, WMI, and two IQ tests (Wechsler´s IQ Test and Cattell´s Culture Fair Test) were employed.

Subjects. In total 181 German subjects served in both experiments. Administered in single subject sessions a booklet in German language presented each stimulus and the scale on one page. The mean IQs were found to be clearly higher than average (M=116, SD=12 for the Wechsler Test and M=113, SD=12 for the Cattell test). The ages ranged from 9 years to 27 years. The subjects were grouped according to the age limits in the German criminal law, below 14 years (N=72, 44 female, M=11.4, SD=1.4), from 14 to 20 years (N=69, 40 female, M=17.4, SD=1.8), and above 20 years (N=40, 29 female, M=22.3, SD=1.3).

Preliminary analyses. There were no interactions of gender or IQ in the following results on the mean judgments. Also, one may note that this kind of work is not like the standard experiment where the main effect is in doubt. Instead, its concern is to reveal patterns of integration in moral judgment. In particular, the 3 x 3 design intends to test the averaging rule by comparing the effects of the two-factorial and one-factorial stimuli. Thus, it may suffice to say that all main effects of the following report were statistically significant at least at alpha=0.001 except as otherwise indicated.

RESULTS

INTEGRATION OF SOCIETAL RISK AND PERSONAL RISK

The Heinz experiment. The main result of the Heinz experiment supports the averaging model. This is shown in the dashed curve of Figure 1 for the one-factorial stimuli of Societal Risk. The dashed curve crosses the curve for the medium level of Personal Risk. Thus, the medium level of Personal Risk either averages up the low level of Societal Risk, or averages down the high level of Societal Risk.

Additionally, the separated points on the right for the one-factorial stimuli of Personal Risk point to the same crossover result. Their spread is larger than the spread of the medium level for Societal Risk in the solid curves and enclosed the latter. The statistical analyses showed highly significant support for both cross-overs: $F(2,254)=14.71$ for Personal Risk and 20.60 for Societal Risk ($p<.001$).

The near-parallelism of the upper and lower curves in Figure 1 was supported statistically ($F(2,360)<1.00$) and implies an additive integration of the two moral informers which validates the linearity of the response. However, this sign of a "constant weight" moral algebra was not supported due to a significant Personal Risk x Societal Risk interaction: $F(4,720)=6.75$, $p<.001$. Figure 1 shows the cause of that non-additivity. The curve for the medium level of Personal Risk had a flatter slope than the curves of the two extreme levels. That deviation of the middle curve from parallelism is in line with the differential weight averaging model. It supports a smaller weight of medium Societal Risk than of low or high Societal Risk.

Another important result is that both stimulus variables, Societal Risk and Personal Risk, had strong effects on judgments of deserved punishment. This is visible in the slopes of each solid curve and in the vertical distances between the three curves in Figure 1. The slopes represent the effects of the three levels of Societal Risk and the distances represent the effect of the Personal Risk variable. The impact of both variables supports the operation of integration.

Age did not affect the interaction of single and combined stimuli, the interaction of Personal Risk and Societal Risk, nor the main effects. This was in contrast to an age effect found by Colby, Kohlberg et al. (1983). In particular, they found that with the "standard issue scoring" of the MJI the frequencies of stage I decreased from 29 % to 0 % within the ages of 10 years to 14 years, whereas the frequencies of stage IV increased from 0 % to 30 % from 10 to 21 years.
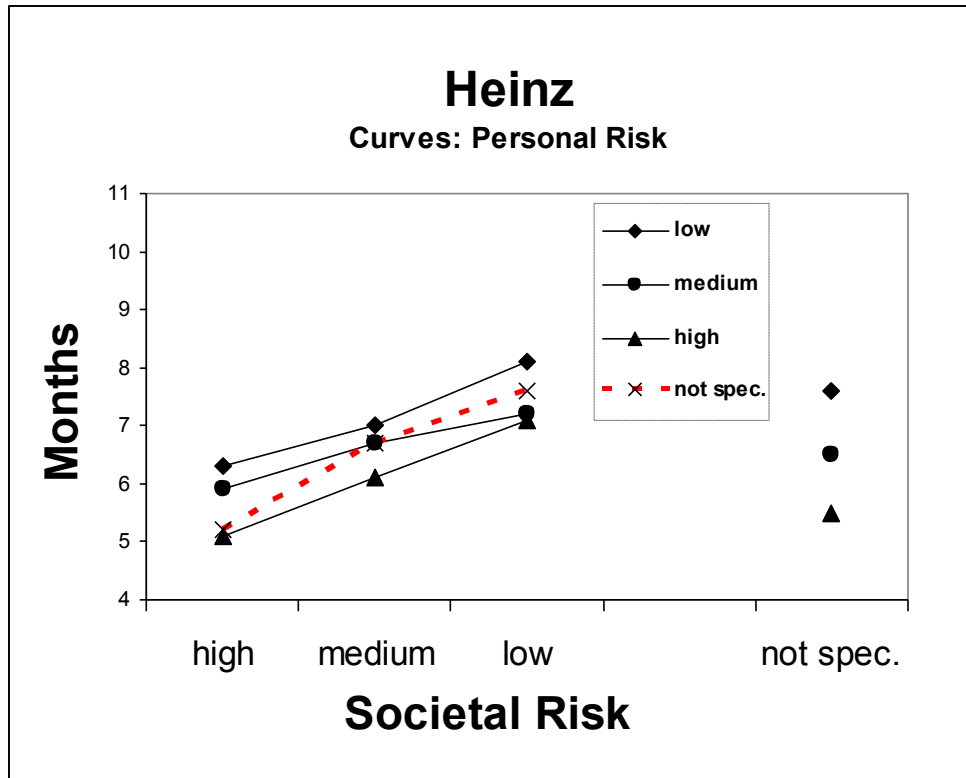
Figure 1: The mean punishments of the Heinz experiment as a function of the three levels of Societal Risk (horizontal axis) and the three Levels of Personal Risk (curve parameters). The graph shows graphical support for the averaging model of information integration theory by the cross-over of the dashed curve with the middle curve.

Since in MJI research stage 1 was found only below the age of 14 and 60% of the sample were older, the large effect of Personal Risk may contradict an expectation of the Kohlbergian approach in which Personal Risk as the informer of the first stage should have a much lower impact on the ratings than Societal Risk as the informer of the fourth stage. However, there was a correspondence of both approaches. Societal Risk had a larger effect than Personal Risk as measured by the upper and lower curves. The effect of Societal Risk is 1.7 months as shown by the slopes of the upper and lower curves of Societal Risk. The effect of Personal Risk is 1.0 months as visible in the spread of the curves. Thus, the effect of Societal Risk was almost 2 times larger than the effect of Personal Risk.

Note that the low risk conditions got a higher punishment rating than the medium or high risk conditions for both variables. This shows that the subjects did not mechanically react to the stimulus levels by giving higher punishments for higher stimulus levels. The validity of this pattern was supported by a collateral result from the training phase where the high conditions of Personal Risk and Societal Risk were chosen as morally demanded thought by 97 % (Personal Risk) or 98 % (Societal Risk) of the subjects.

The Paul experiment. The results for the Paul experiment are very similar to those of the Heinz experiment.

Again, the main result for the Paul experiment supports the averaging model as the dashed curve of Figure 2 for the one-factorial stimuli of Societal Risk crossed the curve for the medium level of Personal Risk. Additionally the separated points on the right for the one-factorial stimuli of Personal Risk is larger spread than the spread of the medium level for Societal Risk in the solid curves and enclosed the latter. The statistical analyses showed highly significant support for both crossovers: $F(2,254)=9.22$ for Personal Risk and 6.49 for Societal Risk ($p<.001$).

As in the Heinz experiment both stimulus variables, Societal Risk and Personal Risk, had strong effects on judgments of deserved punishment as shown by the slopes of each solid curve for Societal Risk and by the vertical distances between the three curves in for the Personal Risk variable.

Again, the statistically supported ($F(2,360)<1.00$) near-parallelism of the upper and lower curve in Figure 2 implies an additive integration of the two moral informers and again this sign of an additive moral algebra was not supported by a significant Personal Risk x Societal Risk interaction: $F(4,720)=4.41$, $p=.002$. Again, Figure 2 shows that the curve for the medium level of Personal Risk has a flatter slope than the curves of the two extreme levels. As noted before the deviation of the middle curve from parallelism is in line with the differential weight averaging model. It supports a smaller weight for medium Societal Risk than for low or high Societal Risk.

Note that the effects of both informers are smaller in the Paul experiment than in the Heinz experiment. However, again Societal Risk had a lager effect than Personal Risk as measured by the upper and lower curves. The effect of Societal Risk is 1.3 months as shown by the slopes of the upper and lower curves of Figure 1. The effect of Personal Risk is 0.6 months as shown by the spread of the curves. Thus, the effect of Societal Risk again was around 2 times larger than the effect of Personal Risk.
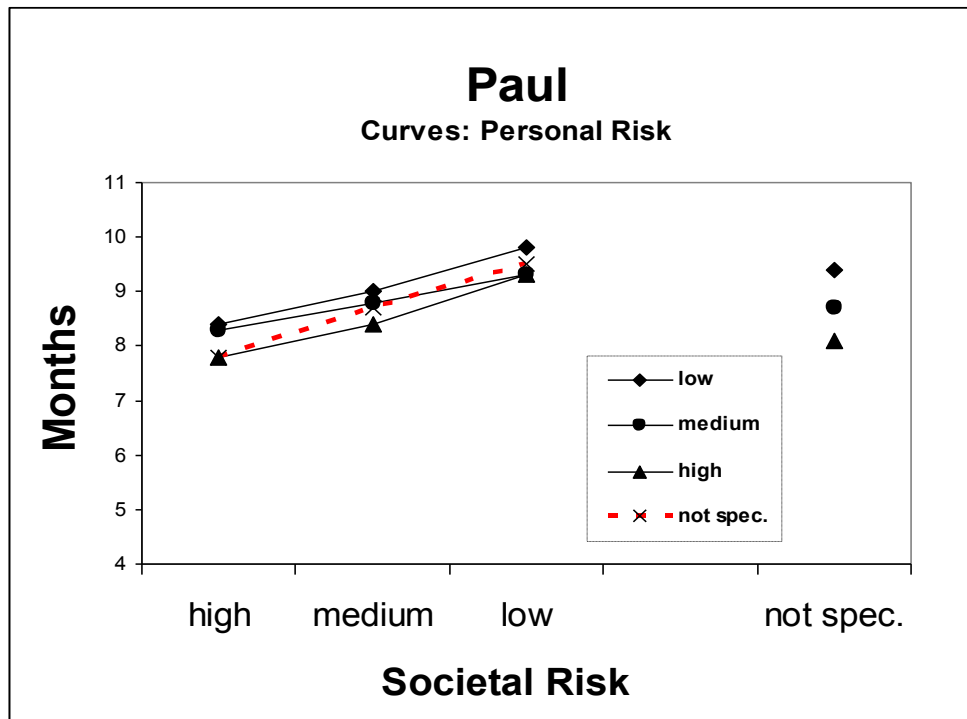
Paul
Curves: Personal Risk

Figure 2: The mean punishments of the Paul experiment as a function of the three levels of Societal Risk (horizontal axis) and the three Levels of Personal Risk (curve parameters). The graph shows graphical support for the averaging model of information integration theory by the cross-over of the dashed curve with the middle curve.

As with the Heinz experiment, age did not affect the interaction of single and combined stimuli, the interaction of Personal Risk and Societal Risk, and the main effects on the punishment ratings. Moreover, the low risk conditions got a higher punishment rating than the medium or high risk conditions for both variables showing again that the subjects did not mechanically react to the stimulus levels by giving higher punishments for higher stimulus levels. The validity of this pattern was supported as well by the collateral result from the training phase where the high conditions of Personal Risk and Societal Risk were chosen as morally demanded thought by 96 % (Personal Risk) or 96% (Societal Risk) of the subjects.

Initial judgments and training phase. The mean initial judgments for the Heinz dilemma were 5.7, 4.9, and 3.7 months imprisonment for the three age levels, $F(2,178)=6.24$, $p=0.002$), and were larger and of opposite

age effect for the Paul dilemma, 8.2, 9.4, and 10.1 months, $F(2,178)=11.48$, $p<0.001$. On average, age increased the punishment for stimuli with thoughts of Paul, 7.7, 9.2, 10.1 months, but did not for stimuli with thoughts of Heinz, 6.5, 6.5, 6.6 months, for the three age levels respectively. Apparently punishment for Paul´s burglary is higher than for Heinz´. Notably, harsher subjects as measured by their initial judgment were in general harsher than lenient subjects. However there were no different patterns in the results presented above concerning the effects of Personal Risk and Societal Risk as covariance analyses showed.

Thoughts about the "window exposition" as an informer had no effect on the punishment ratings. Correspondingly, "window exposition" was chosen as a morally demanded thought by only 2 % (Heinz) or 3 % (Paul) of the subjects. Therefore, thoughts about the "window exposition" appeared as a morally irrelevant informer supporting that the effects of Personal Risk and Societal Risk show their moral relevance.

The informer "friends would have (have not) acted the same way" representing Kohlberg´s stage III had an effect of 0.8 months on the punishments ($F(1,178)=26.49$, $p<.001$). The "friends" effect was clearly smaller in comparison to Societal Risk when presented only (2.4 or 1.7 months for Heinz or Paul respectively). Note, that in the MJI the frequency of the related Kohlbergian stage III was near 50 % at the age of 21 years, which is larger than the frequency of stage IV (Colby, Kohlberg, et al., 1983, Figure 1, p. 46). Contrarily to that extremely higher frequency of stage-III-reasoning in the MJI the effect of the "friends" informer was even lower than that of Personal Risk when presented only (2.1 or 1.3 months for Heinz or Paul). But similarly to Kohlberg´s MJI results where stage III frequencies slightly decreased within the age range of the present experiment, the effects for the three age groups were 1.0, 0.6, 0.7 months in the Heinz experiment, and 1.0, 0.9, 0.3 months respectively in the Paul experiment.

Correspondingly to their low effect regarding the punishment ratings, the thought "friends would do so" was chosen only by 8 % (Heinz) or 4 % (Paul) of the subjects as morally demanded thought in the training phase. Therefore, the actor´s thoughts about what friends think appear as a weak moral informer in the novel thought task. This result may need more consideration in future research.

INDIVIDUAL DIFFERENCES

Individual differences are of major concern in moral judgment. Two approaches are employed: Correlational interdependencies of individual effect sizes and cluster analyses with the nine judgments on combined stimuli (see Hofmans & Mullet, in press).

Correlational interdependencies. Individual difference variables were calculated by subtracting the (lower) punishments for the high risk levels from the (higher) punishments for the low risk levels.

Kohlberg´s stage sequence predicts that those individual differences of the levels of different Kohlbergian moral stage informers should correlate negatively. This can be derived from the MJI data which Colby, Kohlberg, et al. (1983, pp. 48 – 49, Figures 2 to 5) reported about four subjects of their longitudinal research. Similarly to the results of Hommers & Lee (2010), the MJI hypothesis was rejected. Every correlation among the effects of stage I and stage IV informers was positive ($p<.001$) and varied between $r=0.52$ and $r=0.80$.

Among the effect correlations of stage I and stage IV informers with the stage III informer all correlations within the Paul experiment and two within the Heinz experiment were positive. However, these correlations were lower than that among the stage I and stage IV informers as they varied between $r=0.10$ and $r=0.46$.

The individual effects of the non-moral information should be uncorrelated with the individual effects of the Kohlbergian stage informer. This prediction was well supported as the correlations were near zero and smaller than those among the Kohlbergian differences, varying between $r=-0.17$ and $r=0.13$. One negative correlation even reached significance at the 5%-level.

Clusters. The patterns of the three clusters obtained with the Kmeans approach were more or less similar to the patterns of Figure 1 and Figure 2. Further graphical and/or statistical support for the averaging model could be obtained as the judgments on single stimuli had larger effects than those of the combined or showed graphical crossovers as in Figure 1 and Figure 2. In addition the non-additivity of integration and the lower slope of the medium level of Societal Risk appeared in those three clusters, too.

The clusters, which showed no association to age, gender, and WMI, may be characterized as:

- Strong Reactors (*N*=58 in the Heinz experiment and *N*=15 in the Paul

experiment): Their graphs showed larger slopes and spreads than the graphs of the other clusters (3.3 and 4.9 months for the effects of Societal Risk as well as 2.4 and 2.5 months for those of Personal Risk for Heinz and Paul respectively). Their initial judgments about the vignettes alone (being not included in the classification) were 6.5 and 9.7 months respectively;

- Low Punishers (N=39 in the Heinz experiment and N=28 in the Paul experiment): Their mean punishments were between two and six months (with small effects 1.0, 1.4 months for Societal Risk and 0.7, 0.7 months for Personal Risk for Heinz and Paul respectively). Their initial judgments about the vignettes alone were 2.8 and 5.8 months respectively;

- High Punishers (N=52 in the Heinz experiment and N=99 in the Paul experiment): Their mean punishments were between seven and twelve months (also with small effects 1.8, 1.1 months for Societal Risk and 1.0, 0.5 months for Personal Risk for Heinz and Paul respectively). Their initial judgments about the vignettes alone were 6.5 and 10.4 months respectively.

A fourth cluster of No Reactors was found with around one sixth of the subjects and was irregularly showing essentially no stimulus effects. In terms of information integration, those subjects may have found the moral informers not important enough to consider them in their judgments. Their initial judgments on each vignette alone were like generally observed, higher for Paul (N=39) than for Heinz (N=32), 5.3 and 7.9 months respectively for Heinz and Paul.

TRADITIONAL NON-INTERVIEW

Three groups with high (N=53, age M= 13, SD= 4.0), medium (N=66, age M= 16, SD= 3.8), and low (N=62, age M= 18.5, SD= 4.2) WMI-scores for arguments on stage III and stage IV were analyzed.

Most notably, with all three WMI groups the crossovers of single and combined presentations were found in the Heinz experiment as well as in the Paul experiment, supporting independence of the average model from other quantitative measures of moral judgment.

For the effect sizes two hypotheses can be stated. First, there should be a larger effect of the moral informer related to stage IV in subjects with a higher WMI-score. Second, there should be a smaller effect of the moral informer related to stage I in subjects with a higher WMI-score.

As expected by the Kohlbergian model, the effect of Societal Risk increased significantly with the WMI-level for Heinz and for Paul. As Table 1 shows in the two Societal Risk columns, the effects of Societal Risk were 1.9, 2.1, 3.3 months (Heinz) as well as 1.4, 1.0, 2.8 months (Paul) for the single presentation, and 1.1, 1.5, 2.4 months (Heinz) as well as 1.0, 1.0, 2.8

months (Paul) for the combined stimuli in the two experiments for increasing WMI respectively.

Table 1: Effect Sizes in the two experiments for three groups of WMI-scores* and for the total group.

Risk Informer	HEINZ Experiment				PAUL Experiment			
	Personal		Societal		Personal		Societal	
Design	Combi	Single	Combi	Single	Combi	Single	Combi	Single
WMI: Below 43 (N=72)	0.6	1.5	1.1	1.9	0.6	1.1	1.0	1.4
WMI: 43 to 46 (N=69)	1.0	1.9	1.5	2.1	0.5	0.9	1.0	1.0
WMI: Above 46 (N=40)	1.4	2.8	2.4	3.3	0.8	1.8	1.8	2.8
Total	1.0	2.1	1.7	2.4	0.6	1.3	1.3	1.7

*: Note that the effect of Friends´ thought decreased in the training phase, 1.5, 0.3, 0.7 months for Heinz and 1.3, 0.7, 0.6 months for Paul (for the three WMI levels respectively), being marginally significant, $F(2,178)=2.37$ (p=.096) in contrast to expectations from Kohlbergian MJI results.

But contrarily to expectation from Kohlberg´s model, the effect of Personal Risk also increased with WMI-level (significantly for Heinz or non-significantly for Paul). This is shown in the Personal Risk columns of Table 1. Those effects were 1.5, 1.9, 2.8 months (Heinz) as well as 1.1, 0.9, 1.8 months (Paul) with the single presentation, and 0.6, 1.0, 1.4 months (Heinz) as well as 0.6, 0.5, 0.8 months (Paul) in combination with Societal Risk in the two experiments for increasing WMI respectively.

The stimulus effects of Table 1 imply that the ratios of the effects from single and combined stimuli differed among the WMI groups and were larger for Personal Risk (2.50, 1.90, 2.00 for Heinz and 1.83, 1.80, 2.25 for Paul, for the three WMI groups respectively) than for Societal Risk (1.73, 1.40, 1.38 for Heinz and 1.40, 1.00, 1.56 for Paul). The general implication is that the weight of Societal Risk is larger than the weight of Personal Risk in all WMI groups because the addition of the Societal Risk informer in the combined stimuli reduces the effect of the Personal Risk informer more than vice versa.

The averaging model allows for estimating the relative sizes of the weights for Personal Risk and Societal Risk from the quotients of effect ratios "single/combined". Assuming that the weight of the initial state is constant it follows from the algebraic equation of the averaging model that the weight of Societal Risk is at least around one and a half greater (1.5) than the weight of Personal Risk, with small variations among the WMI-groups.

DISCUSSION

The present results give profound answers to several questions. First, the unification of two major approaches to moral judgment is feasible. Typical content of the Kohlbergian stages can be used as stimulus for integration studies. Second, the form of the integration process of two Kohlbergian moral informers, Personal Risk and Societal Risk, could be identified. The use of a proper integration design showed supported by the cross-over test that those moral informers are integrated according to an averaging model with one incident supporting differential weighting. Third, the generality of the information integration of the two employed Kohlbergian moral informers with regard to situational and individual variables was shown. The support of generality was observed (a) for the two scenarios of Heinz and Paul telling about two actors with altruistic versus egoistic motives, (b) for three groups with different levels of another non-interview measure closely related to the Kohlbergian stage concepts, (c) for three clusters presenting around five sixth of the subjects, and (d) by independence of those results from harshness, age, and gender of the subjects.

Unification. The results demonstrate that Kohlbergian stage theory can be unified with the theory of information integration by using Kohlbergian concepts within the methods of functional measurement. Indeed, even the two Kohlbergian stage concepts of stage I and stage IV

being rather separated in the stage sequence are still operative as multiple determinants of moral judgment and may consequently be subject to some form of information integration. Discrepancies and confirmations to expectations based on Kohlberg´s theory should not distract from the main point that both approaches can join to enhance understanding of morality.

As a consequence of unification unproven claims of stage theorists which are not empirically approachable by the stage interview approach (Anderson, 2008) have been confirmed by the present information integration study. For example the hypothetic view of Colby, Kohlberg et al. (1987, p. 2) who assumed that moral concepts are not „used independently of one another but rather are bound together by common structural features" and that their focus is "on the relations among ideas in the individual´s thinking" became visible in the integration pattern for Personal and Societal Risk. The obtained support for integration also demonstrated that Kohlbergian stage concepts are "involving the integration of the various considerations" (Rest 1983, p. 561). But the stages do not provide correct "frameworks for prioritizing and integrating considerations" (Rest, 1983, p. 563). Instead, those frameworks turned out to be incidents of information integration like those found with the integration of moral informers in other areas of morality, e.g. Piagetian centration theory of moral realism (among others Leon, 1980; Surber, 1982; see Anderson, 1991, chapter 5; 1996, chapter 7; 2008, chapter 7 for further details). By categorizing interview statements these explications of the integration process may not be possible in Kohlberg´s theory, although the importance of integrating multiple determinants was recognized when fairness was equated with "balancing or weighing of conflicting claims" (Colby, Kohlberg et al., 1983, p. 7). Instead of categorizing interview statements into stages the natural laws of moral information processing become visible by direct inquiry with the methods of information integration theory.

As shown, the operation of information processing becomes visible when typical moral contents are used as stimuli for eliciting punishment judgments. This has been done here with the examples of the two well separated stages heteronomous morality (stage I) and social system and conscience (stage IV) which were chosen for several reasons. First, Personal Risk and Societal Risk appear to be relevant aspects of judgments about criminal behavior independently of specific moral theories. Second, although Kohlberg´s stage theory assumes currently five stages (Colby, Kohlberg et al., 1987) ranging from the morality of obedience through the morality of law and duty to post-conventional morality, the postconventional stage V was found rarely. Third, Gibb´s two-phase model (Gibbs, 2003; Gibbs et al., 2007) cancelled even the fifth stage and used

only the first four of Kohlberg´s six stages which were the original two preconventional and two conventional stages. Fourth, according to a widespread basic agreement in the literature about the presence of the first four stages it appeared most effective for testing the integration of Kohlbergian informers to use the extremes of those first four stages.

The demonstrated integration of those two generally agreed Kohlbergian stage contents fits well to reports that showed integration of other Kohlbergian concepts and of Kohlbergian and non-Kohlbergian moral concepts like recompense (Hommers, 1997; Hommers & Lee, 2010). Apparently, the IIT approach resolves the problem that the hermeneutical procedure and the theoretical basis of sociomoral perspectives include no theory about balancing or weighing of conflicting claims and of other important moral concepts. Therefore, the present study confirms the claim of unification of two major approaches to moral judgment.

Moral Algebra. Strong support for the notion of moral algebra was obtained. The idea of moral algebra goes back to Aristotle´s model of distributive justice but recent information integration research has been shown moral algebra operative in other topics (see Anderson, 1991, chapter 5; 1996, chapter 7; 2008, chapter 7 for details). The present confirmation of the averaging model of the moral algebra of Personal Risk and Societal Risk fits well to other evidence for moral algebra, e.g. in the blame scheme of intent and consequences or in equity research (see Anderson, 1991).

The integration process was identified as an averaging one due to the strong crossover test of partial and complete stimuli. The parallelism found in the graphs for the extreme levels (high and low) of Societal Risk at least support the response linearity. But, there was a non-additive component in the results as supported consistently by the statistically significant interaction and by one non-parallel curve in the graphs of the combined judgments for the medium level of Societal Risk. The concept of differential weights in IIT allows for this non-additivity.

Because sensitivity of the punishment response to age and scenario was shown in the results, the invariance results for the differentially weighted averaging are noteworthy. The differential weighting result was independent of the altruistic/egoistic content (Heinz/Paul), was independent of personality variables like age, gender, harshness, was independent of the classification into three regular clusters and into the three WMI groups which served as another approach to a non- interview assessment of moral judgment.

Usefulness of the personal thought scenario. The simple novel task of presenting thoughts of a harmdoer has several advantages. Methodological problems of Kohlberg´s task are avoided: (1) extensive training for the lengthy and individually administered verbal protocol of the Moral Judgement Interview, (2) mistakes due to the subjectivity of the person who scores the verbal protocols by a scoring manual, (3) any implicit assumption of the interpretative stage diagnosis from verbal protocols (Colby, Kohlberg, et al., 1987), (4) the objection that the standard protocol method confounds moral stage with verbal ability.

The less demanding response and the experimental control of the stimulus situation can be employed with elementary school children and in cross-cultural research (Hommers & Lee, 2010) to study a major period of moral development at preschool and elementary school ages. Finally, the novel task has ecological validity as it proceeds like the interrogation of a judge who is interested in the motives of the culprit and who might simply ask, "what did you think when you were violating the law?" By putting the subjects into the judge´s place they can reveal their moral capacities.

REFERENCES

Anderson, N.H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology, 70,* 394-400.

Anderson, N.H. (1981). *Foundations of information integration theory.* New York: Academic Press.

Anderson, N.H. (1991). Moral-social development. In N.H. Anderson (Ed.), *Contributions of Information Integration Theory. Vol. III: Developmental* (pp.136-187). Hillsdale, NJ: Erlbaum.

Anderson, N.H. (1996). *A Functional Theory of Cognition.* Mahwah, NJ: Erlbaum.

Anderson, N.H. (2008). *Unified Social Cognition.* New York: Psychology Press.

Anderson, N.H., & Butzin, C.A. (1978). Integration theory applied to children´s judgments of equity. *Developmental Psychology, 14,* 593-606.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development, 48,* 1-94.

Colby, A., Kohlberg, L., Speicher, B., Hewer, A., Candee, D. Gibbs, J., & Power, C. (1987). *The measurement of moral judgment. (Vol. I, Vol. II).* New York: Cambridge University Press.

Ehrmann, D. (2010). *Validitätsprüfung des Richterspiels.* Unpublished Diplomathesis. Universität Würzburg: Institut für Psychologie.

Gibbs, J. C. (2003). *Moral development and reality: Beyond the theories of Kohlberg and Hoffman.* Thousand Oaks, CA: Sage.

Gibbs, J.C., Basinger, K.S., Grime, R.L., & Snarey, J.R. (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review, 27,* 443–500.

Hofmans, J., & Mullet, E. (in press). Towards unveiling individual differences in different stages of information processing: A clustering-based approach. *Quality & Quantity.*

Hommers, W. (1997). Integration of Kohlbergian information in punishment. *European Journal of Applied Psychologie (Revue Européenne der Psychologie Appliqué), 47*, 31-37.

Hommers, W., & Lee, W.Y. (2010). Unifying Kohlberg with information integration: The moral algebra of recompense and of Kohlbergian moral informers. *Psicológica, 31*, 689-706.

Hommers, W., & Lewand, M. (2005). Zum Einfluss des Tatmotivs auf eine Voraussetzung der strafrechtlichen Verantwortungsreife. *Monatsschrift für Kriminologie und Strafrechtsreform, 88*, 61-69.

Kohlberg, L. (1976). Moral stages and moralization: the cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behaviour: Theory, research, and social issues* (pp. 31-53). New York: Holt, Rinehart & Winston.

Leon, M. (1980). Integration of intent and consequence information in children´s moral judgment. In F. Wilkening, J. Becker, & T. Trabasso (Eds.). *Information integration in children* (pp.71-91). Hillsdale, NJ: Erlbaum.

Leon, M. (1982). Rules in children´s moral judgments: Integration of intent, damage, and rationale information. *Developmental Psychology, 18*, 835-842.

Leon, M. (1984). Rules mother and sons use to integrate intent and damage information in their moral judgments. *Child Development, 55*, 2106-2113.

Lewand, M., & Hommers, W. (2001). *Würzburger Moral Inventar*. Würzburg, Institut für Psychologie (unpublished manuscript).

Modgil, S., & Modgil, C. (1986). *Lawrence Kohlberg. Consensus and Controversy*. Philadelphia: Falmer Press.

Piaget, J. (1965). *The moral judgment of the child*. (M. Gabin, Trans.). New York: Free Press (original work published 1932).

Rest, J.R. (1975). Longitudinal study of the Defining Issues Test of moral judgment: A strategy for analyzing developmental change. *Developmental Psychology, 11*, 738-748.

Rest, J.R. (1983). Morality. In P.H. Mussen, J.H. Flavell, & E.M. Markmann (Eds.), *Handbook of child psychology: Vol. III. Cognitive development* (4th. ed., pp 556-629). New York: Wiley.

Rest, J.R. (1986). Moral research methodology. In S. Modgil, & C. Modgil (Ed.), *Lawrence Kohlberg. Consensus and Controversy* (pp.455-469). Philadelphia: Falmer Press.

Rest, J., Narvaez, D., Bebeau, M.J., & Thoma, S.J. (1999). *Postconventional Moral Thinking. A Neo-Kohlbergian Approach*. Mahwah, NJ: Erlbaum.

Surber, C.F. (1982). Separable effect of motives, consequences, and presentation order on children´s moral judgments. *Developmental Psychology, 18*, 257-266.