

A Cognitive-Representational Account of Intuitive Moral Judgment: Effects of Typicality and Accessibility

Niklas Fransson* and Karl Ask

Department of psychology, University of Gothenburg, Sweden

Abstract: In this article, it is argued that intuitive judgments of immoral events result from an automatic process where perceived events are matched against mentally represented event prototypes. The proposed cognitive underpinnings of such a process are tested in two experiments. Experiment 1 demonstrated that typical immoral events require shorter judgment times than atypical events. This typicality effect implies that immediate moral responding depends on the similarity of an encountered event to a pre-existing mental prototype. Experiment 2 showed that priming representations of immoral events facilitates the responding only to other events violating the same moral value, and not to events related to other moral values. This finding provides further support for the notion that moral reactions rely on pre-existing schematic mental representations, and suggests that these representations are stored in associative networks with values as a basis for categorization. It is concluded that the results concord with and extend recent work that places moral cognition in a dual-process perspective.

Keywords: Moral judgment, Dual process, Cognitive structure, Response time.

INTRODUCTION

Traditionally, values have been considered an important factor guiding moral judgment and action [1-4]. Values are commonly considered as “desirable transsituational goals, varying in importance, that serve as guiding principles in the life of a person or other social entity” [5(p21)]. This means that they are general cognitive structures, applicable to a wide range of situations, rather than confined to specific instances. The influence exerted by a value on behavior is described by Feather [6] who argues that the value is used as a frame of reference when evaluating the attractiveness or desirability of different objects or acts. More specifically, when encountering a novel stimulus, the implications of the object at hand or potential actions are compared to the desired outcome as dictated by the value. If they are in agreement, the stimulus or act is perceived as favorable, whereas disagreement leads to an aversive interpretation. This evaluation, in turn, informs the perceiver about how to respond to the specific instance, and whether to adopt a positive or negative attitude [6, 7].

Traditional theories in moral psychology have generally been concerned with the function and development of morality [8-10]. That is, a stage-like approach has been constructed to explain how experience and socialization lead to mature moral reasoning. However, although values are acknowledged as important elements in this process, these approaches say nothing about the cognitive structure of moral values. More recent research has been devoted to specifying the structure of values [5, 11-13] as well as the

process by which values are applied to various situations [6, 7]. These later efforts, however, say more about the structural interrelations *between* values and less about the internal cognitive structure of single values. In order to gain a fuller understanding of moral judgments its cognitive underpinnings need to be identified and explored.

Since moral judgment by tradition is considered to be a process of reasoning, most research has focused on reactions to moral dilemmas. However, it could be argued that far from all situations with moral implications take on the characteristics of a dilemma. In fact, many of the moral transgressions that people encounter are easily detected and obvious, and we react to them intuitively. In the present study the main focus is on the cognitive structure underlying these intuitive reactions. We argue that moral cognition can be placed in a dual-process perspective; that is, moral reactions can be viewed as results of more or less controlled or automatic processes [e.g., 14, 15]. Whereas conscious elaboration at the controlled end of the continuum is probably responsible for judgments in unfamiliar moral *dilemmas*, processes at the automatic end are more likely to mediate *spontaneous* moral reactions to obvious transgressions. That is, in general terms we propose that the process of moral reaction and judgment is automatic and fast when the event/stimuli characteristics are readily recognized as a moral transgression, while novelty among event attributes prompts slower, more elaborated processing. A further assumption of our model is that increased familiarity with an event entails an automatization process [16]. That is, as a novel event is encountered repeatedly and its attribute elaborated upon, the perceiver gradually develops a memory representation of the event. Once the representation has achieved sufficient stability as a concept, the moral response that was formerly pro-

*Address correspondence to this author at the Department of psychology, University of Gothenburg, Sweden; Tel +46 31 786 1684; Fax; +46 31 786 4628; E-mail: niklas.fransson@psy.gu.se

duced by deliberate reasoning can be triggered automatically through pattern matching. Conscious deliberation thus precedes automatic processes by establishing associative links in memory between the perception of an event and the typical response to the event. Once automatized, the immediate moral reactions to familiar events possess all four qualities of automatic processes identified by Bargh [17]; that is, the perceiver is unaware of the judgment process, it proceeds rapidly and efficiently, it does not require the perceiver's intention, and it occurs mainly without the perceiver's control. The proposed relationship between the automatic and deliberate routes to moral judgment is illustrated in Fig. (1).

The characterization of morality based on dual cognitive processes is not new. During the past decade attempts to incorporate the logic of dual-process models have been made by moral psychologists [18-20]. A prerequisite for automatic processing of moral stimuli, however, is the existence of mental representations that can trigger a response to an event without the need for elaboration. To date, no systematic investigation of these representations has been conducted.

In an attempt to provide a more detailed cognitive framework, Biel, Fransson and Dahlstrand [21] put forth a

structural model of moral values. The model is based on the notion of theory-based, as opposed to similarity-based, concepts [22, 23]. Briefly, the theory-based view holds that knowledge is the organizing principle of concepts, and that properties of concepts exist on different levels of abstraction. That is, instances are grouped into categories based on some underlying meaningful commonality. In contrast, similarity-based concepts cohere because of shared superficial characteristics, typically on a perceptual level. Although contrasting similarity-based and theory-based concepts, Medin [22, 23] proposes that identification of instances as members of either type of category involves attribute matching (similarity) as an integral part. In accordance with this proposal, the model of Biel *et al.* adopts a view of moral values as concepts, with information pertinent to these concepts cohering because of its shared moral implications. In addition, the conceptual structure of each moral value is hierarchically organized, represented by norms and schemas (see Fig. 2).

At the most abstract level there is a norm (e.g., "One should be true and sincere with persons close to oneself"), which represents proper moral conduct with respect to the value. Norms thus constitute the core meaning around which the concept is organized. At a more specific level, schemas

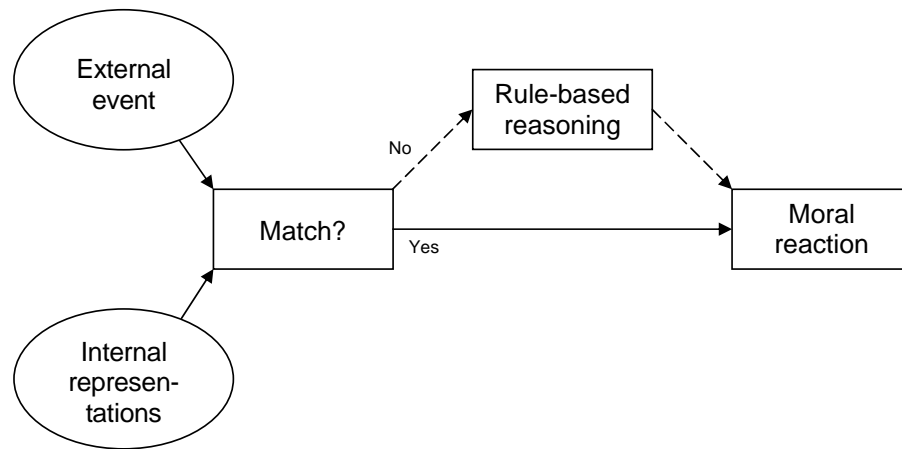


Fig. (1). Schematic representation of the hypothesized automatic (continuous line) and deliberate (dashed lines) routes to moral reactions.

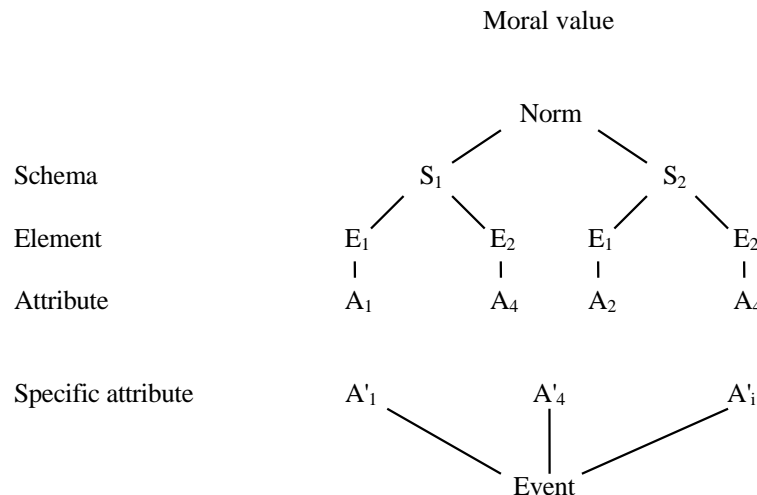


Fig. (2). Components in a moral value matched against a hypothetical event.

represent events in which the norm is violated. The features of the schemas are represented at two levels of abstraction: 'elements', which represent core components of events (e.g., act, intention, victim), and 'attributes' which are the specific instantiations of the elements (e.g., lying, personal gain, friend). Schemas thus correspond to Medin's conception of information on a more perceptual level of a concept (i.e., attributes) than that represented in norms, and allow for quick recognition of immoral events. It is proposed that the activation of a moral value depends on the match between the features of an encountered situation and the attributes of a schema. A match signals a norm violation, activates the value, and triggers a moral reaction. This reasoning is similar, although differing in terminology, to that of Mandler [24] who holds that values are represented by schemas at different levels of abstraction, and that the activation of a schema is automatic when an encountered event is congruent with a schema.

The model was tested by Biel *et al.* [21] who had participants describe the contents of sampled real-life immoral events, using one or a few morally connoted words. By analyzing clusters of words used to describe different events eight distinct categories of immoral events were identified. It was assumed that events in the same category activated the same moral value. In a second study [25], it was found that the salience of different event elements varies between representations of different values. Thus, the relative importance of different elements seems to be crucial to the distinction between different values.

THE PRESENT STUDY

Closely related to the schema concept is the idea of prototypes, which are generally defined as schematic knowledge structures representing the central tendency of a category [e.g., 24, 26]. That is, properties of instances within a category are abstracted by the perceiver through repeated encounters, resulting in a 'mean representation' incorporating the most common features shared by category members. Consequently, the most frequently encountered member makes the greatest contribution to the category prototype. The identification of an instance as a category member is considered to depend on its similarity to the prototype, through a process that closely resembles the attribute matching proposed in the model of Biel *et al.* [21]. Research on cognitive event representation has demonstrated that social events are often represented as prototypes [27-29]. Thus, it seems plausible that immoral event representations are prototypically organized as well. The present study was conducted in order to test this assumption empirically.

EXPERIMENT 1

If immoral events are stored as prototypes in memory, a typicality effect in event classification is to be expected [30, 31]. That is, events highly similar to an event prototype should be classified as immoral more immediately than non-prototypical events. Stimuli sentences were generated that depicted events that violate a moral norm. Since more frequently encountered instances of a category are assumed to make greater contributions to the category prototype [22, 32], typicality of stimuli events was manipulated by varying

the base rate by which the events were expected to occur in real life. Thus, two sets of events were created, one of which contained events that are frequently occurring (typical), and one containing rare events (atypical). The hypothesis was that typical events would be classified as immoral more readily than atypical events.

METHOD

Participants

A total of 27 undergraduate psychology students at the University of Gothenburg (13 males and 14 females), aged 19 to 42 ($M = 25.2$; $SD = 5.9$), volunteered to participate in the experiment. No compensation was offered for participation.

Material

Short target sentences describing immoral events were generated, using the eight categories of immoral events identified by Biel *et al.* [21] as a framework (see leftmost column of Table 1). One event assumed to be highly typical of each category was formulated, by combining the attributes previously found to be the most prominent in the category [25]. For example, the typical sentence from category I read "Employer pays men more than women".¹ This was considered to represent the issue of unfair distribution, which is central to the event category. In addition, a less typical event was generated for each category. This was accomplished by replacing the attributes of the typical sentences with less typical attributes, while maintaining the relationships between the attributes. For example, the atypical sentence of category I read "Judge imposes lighter sentences on men than on women". This sentence still describes the unequal treatment of two groups, but its attributes (i.e., judge, sentence) are less commonly associated in the context of unfair distribution. Word count and length was kept fairly constant between the two versions of a category. On average, typical events were described with 5.4 words containing 29.8 characters and 83 syllables, whereas atypical sentences contained 5.3 words, 32.3 characters and 90 syllables. To create an equal number of immoral and neutral events in the stimulus material, 16 sentences depicting irreproachable actions were formulated, which will be referred to as 'distractors'. Taken together, the entire stimulus material consisted of 8 typical targets, 8 atypical targets and 16 distractors. For the norms describing each category and the respective typical and atypical target stimuli, see Table 1.

Procedure

Participants were tested individually. Participants were given written instructions for the experimental task on a computer screen. Following two practice trials, participants proceeded with the actual experiment. The 32 stimuli sentences were presented sequentially, centered on the screen, in randomized order. For each sentence, participants were to indicate whether the depicted act was immoral or not. A response was given by pressing either a red-colored key on the right-hand side of the keyboard (representing 'immoral'), or a green-colored key on the left-hand side of the keyboard

¹The sentences used in the experiment were in Swedish, and are not comparable to the translated examples regarding grammar, wording, and sentence length.

Table 1. Normative Descriptions of the Eight Categories Identified by Biel *et al.* (1997) and the Typical and Atypical Stimulus Sentences Used in Experiment 1

Category	Norm	Stimulus	
		Typical	Atypical
I	Resources and burden should be fairly distributed	Employer pays men more than women	Judge gives men lighter sentences than women
II	One should take responsibility for one's duties	Drunken man drives car home from the pub	Pilot flies passenger plane despite a heart condition
III	One should be true and sincere with persons close to oneself	Father lets his son down	Grandmother lets her grandson down
IV	The value of (human) life should be sacred	Stepfather batters his stepson	Nurse batters disabled person
V	People in exposed positions should be met with empathy	Schoolchildren laugh at mentally retarded	Policemen laugh at confused woman
VI	People should be treated with respect and trust	Manager refuses to employ immigrants	Doorkeeper refuses gypsies to enter
VII	One should not abuse power	Regime silences critical citizens	Principal silences critical pupils
VIII	One should not betray one's values or principles out of sheer greed	Salesman tricks customers into buying defect cars	Man tricks customers into buying forged travel tickets

(representing 'irreproachable'). Participants were instructed to maximize both the speed and the accuracy of their responses. To make sure that participants' visual attention was focused on the stimulus area, each sentence presentation was preceded by a four-second countdown on the screen. The sentence remained on the screen until a response was given, and the time required to respond was recorded for each judgment.

Directly following the response to an immoral event, a new screen appeared with questions about participants' perception of the depicted event. One question concerned the wrongfulness of the event ("How wrongful do you consider the depicted event?"). A second question was intended to capture the perceived real-life frequency of the event ("How often have you heard of or encountered this particular event?"). Answers to both questions were indicated on 9-point Likert scales. In addition, a third question asked participants whether they had previously been involved personally in a situation of the kind depicted in the sentence ("Yes/No").

RESULTS

Omission of Errors and Outliers

The data were first scanned for erroneous responses; that is, initial key-press judgments of intended immoral events as irreproachable. The average error rate was low across participants (1.3%). However, two participants had given three or more erroneous responses and were excluded from all analyses. The response times for the remaining participants' erroneous responses were omitted from the dataset. In addition, to scan the data for potential outliers, the response times for judgments were first standardized within each category.

Across all eight categories, five (1.3%) extreme observations between 2.7 and 13.8 standard deviations above the mean response time were excluded from analyses, which is well within the range recommended by Bargh and Chartrand [33] concerning trimming of response-time data. Since the experiment employed a within-participant design, and analyses hence required an observation in each cell for each participant, a single missing value would mean that the participant would have to be excluded from the entire analysis. Therefore, the reaction times removed as outliers or due to erroneous responding were replaced with the mean of the participant's responses for the remaining items on the same level of the typicality factor. The number of replaced response times was 20 (5% of the final data set).

Manipulation Checks

In order to establish whether the typicality manipulation had indeed yielded events differing in typicality, participants' ratings of event frequency were entered in a 2 (Typicality: typical vs. atypical) \times 8 (Category: 1 through 8) repeated-measures ANOVA. As expected, the main effect of typicality was highly significant, $F(1, 23) = 261.46, p < .001, \eta^2 = .92$. Typical events were seen as significantly more frequent ($M = 6.2, SE = 0.2$) than atypical events ($M = 2.8, SE = 0.2$). The analysis further revealed a significant main effect of category, $F(7, 161) = 16.50, p < .001, \eta^2 = .42$. Bonferroni post-hoc tests showed that typicality ratings for categories 1 through 5 did not differ significantly from each other. However, category 6 received higher typicality ratings than categories 2 and 8 ($ps < .05$), category 7 received higher ratings than all categories except category 6 ($ps < .01$), and category 8 received lower typicality ratings than all other categories ($ps < .05$). The main effects were qualified by a

significant typicality \times category interaction, $F(7, 161) = 12.84, p < .001, \eta^2 = .36$. However, analyses of simple-main effects showed that the effect of typicality was significant within all eight categories, $ps < .05$ (Bonferroni corrected). As further support of the effectiveness of the typicality manipulation, participants reported to have experienced a significantly higher number of typical events ($M = 2.2, SD = 1.3$) than atypical events ($M = 1.0, SD = 1.0$), $t(27) = 4.96, p < .001, d = 1.91$.

No differences were expected between typical and atypical events in terms of perceived wrongfulness. However, a 2 (Typicality: typical vs. atypical) \times 8 (Category: 1 through 8) repeated-measures ANOVA showed that typical targets were considered significantly more wrongful ($M = 8.0, SE = 0.1$) than atypical targets ($M = 7.6, SE = 0.2$), $F(1, 23) = 10.78, p < .01, \eta^2 = .32$. In addition, there was a significant main effect of category, $F(7, 161) = 18.40, p < .001, \eta^2 = .44$. Post-hoc tests revealed significant differences for 15 of the 28 possible pairwise comparisons between category means, $ps < .05$ (Bonferroni). The typicality \times category interaction was also found to be significant, $F(7, 161) = 4.07, p < .001, \eta^2 = .15$. Tests of simple-main effects showed that the typical event was rated as significantly more wrongful than the atypical event only in categories 2 and 3, $ps < .05$ (Bonferroni corrected). This fact was taken into account in subsequent analyses.

Typicality Effects

Participants' response times (transformed to ms per syllable) for the moral judgments were entered in a 2 (Typicality: typical vs. atypical) \times 8 (Category: 1 through 8) repeated-measures ANOVA. The predicted typicality effect, with typical targets requiring shorter response times ($M = 282$ ms, $SE = 19$) than atypical ones ($M = 317$ ms, $SE = 19$), was statistically significant, $F(1, 24) = 13.76, p < .01, \eta^2 = .36$. In addition, there was a significant main effect of category, $F(7, 168) = 20.53, p < .001, \eta^2 = .46$. Bonferroni post-hoc tests revealed that category 4 (events violating the norm "the value of (human) life should be sacred") required the shortest response times, differing significantly from all other categories ($ps < .05$) except category 8. The longest response times were required in categories 1, 2, and 3, which differed significantly from all other categories ($ps < .05$) but not from each other. Categories 5 through 8 did not differ from each other. The typicality \times category interaction was not significant, indicating that the effect of typicality was consistent across all eight categories.

As mentioned, typical and atypical targets differed in terms of perceived wrongfulness. This raises the possibility that the observed differences in response time can be attributed solely to differences in wrongfulness. To explore this possibility, a separate correlation was computed for each participant between response times and wrongfulness ratings for all 16 target events. The resulting coefficients were transformed to Fisher's Z s to allow for significance testing of means. Indeed, the mean correlation between wrongfulness and response time ($M_Z = -.45; SD_Z = 0.37$) differed significantly from zero, $t(23) = -5.90, p < .001, d = 2.46$. Thus, differences between events in terms of wrongfulness were

able to explain significant portions of the variance in response time. More importantly, however, the mean partial correlation between event typicality and response times when controlling for wrongfulness ratings ($M_Z = .07; SD_Z = 0.20$) also proved significant, $t(23) = 1.77, p < .05$ (one-tailed), $d = 0.74$. That is, even when covariation between wrongfulness and response time was taken into account, typicality explained a unique portion of variance.

DISCUSSION

The results of Experiment 1 are consistent with the proposition that moral values are cognitively represented in terms of event prototypes, and that moral reactions are elicited through a recognition-based matching process between prototypes and perceived event characteristics. The typicality manipulation altered the specific attributes of the depicted situations, and the obtained results indicate that components analogous to the attributes of the value representation model proposed by Biel *et al.* [21] determine the similarity of immoral events to an event prototype.

The main hypothesis, that typical events would require shorter judgment times than atypical events, was supported. Supposedly, the situations that people frequently encounter resemble the memory representations more closely than do rare events, since memory representations are gradually formed through repeated encounters. Thus, the typical targets used in the present study probably displayed a greater similarity to their respective prototypes than did atypical targets. As a result, the pattern-matching process between the event and the representation was executed faster for the typical events. Once again, support is given to the idea that it is information at the attribute level that constitutes the matching features of the immoral event prototypes. The rationale for such a conclusion is that only the attributes and not the event structure differed between the typical and atypical events in the experiment. The typicality effect was consistent across all eight categories, despite considerable differences in the nature of the events between categories. This indicates that manipulations at the attribute level of events have a uniform influence on the perception of a wide range of situations, and lends further support to the validity of the proposed cognitive representation.

Although less relevant to the theoretical issue addressed in Experiment 1, there was a substantial correlation between perceived event wrongfulness and response times. This suggests that the attributes present in highly immoral events may be more salient, thus making it more easy to identify its moral implications. Importantly, however, the typicality effect remained significant even after controlling for perceived wrongfulness. The role of wrongfulness in the proposed cognitive architecture of moral values should be subject to further empirical investigation, but is beyond the scope of the present paper.

EXPERIMENT 2

The ideas tested in Experiment 1 can be extended by adopting the principles of knowledge activation suggested by the propositional model of memory organization [34]. Ac-

According to this model, human memory is organized as an associative network. The network represents events as propositions consisting of nodes and links corresponding to concepts and relations among concepts, respectively. This associative structure implies the principle of *spreading activation*. That is, if a given element in the structure is activated, the state of activation spreads to related elements in the network. This precise mechanism is used to explain priming effects and other phenomena involving knowledge activation in human memory [32, 35]. The application of spreading activation to the domain of moral value representation is straightforward. The recognition and categorization of an immoral event is likely to activate related cognitive structures, namely information contained in the mental representation of the violated moral value. In turn, a temporarily increased activation of the mental value representation would facilitate the categorization of other events transgressing the same moral value.

Experiment 2 was set up to test the principle of spreading activation in relation to moral judgments. We adopted roughly the same procedure as in Experiment 1; the experimental task consisted of brief moral judgments of verbally depicted events, and response times were treated as the main dependent variable. Once again, stimuli events were based on the eight categories of immoral events identified by Biel *et al.* [21]. In contrast to Experiment 1, however, two judgments were to be made in immediate succession. It was predicted that the response to the second event would be facilitated when preceded by an event from the same category, as a result of spreading activation within the mental value representation. No such facilitation was expected when the two events represented different categories, or when an immoral event was preceded by an event without moral implications.

METHOD

Participants

Twenty-one undergraduate psychology students at the University of Gothenburg (5 males and 16 females), aged 19 to 55 ($M = 24.5$, $SD = 7.6$), participated in the experiment. Participants received a lottery ticket (approx. € 2.50) as compensation.

Material

Five events were generated for each of Biel *et al.*'s [21] eight categories, in the same manner as in experiment 1, only this time no manipulation of typicality was made. Events from a particular category were similar only in the sense that the same moral value was violated in each. Thus, the reoccurrence of value-laden words and other perceptual similarities were avoided. In total, 40 immoral events were formulated. In addition, 40 filler events without moral implications were generated.

Procedure

Events were combined into pairs to form one of three experimental conditions: Both immoral events from the same category (SAME condition), two immoral events from different categories (DIFFERENT condition), or a filler event followed by an immoral event (FILLER condition). In addition,

two control conditions were constructed: a filler event followed by another filler event, and an immoral event followed by a filler event. Control conditions were included to balance the number of "immoral" and "irreproachable" responses. The design was balanced so that, across participants, each event occurred in all conditions an equal number of times. Events within a pair were presented sequentially on the computer screen with a 2-second interval between the response to the first event and the presentation of the second event. As in Experiment 1, the task was to judge as quickly as possible whether the event was immoral or not, and to respond by pressing a corresponding red or green key. Each pair of events was separated from the next pair by a distraction task, consisting of a scrambled sentence to be unscrambled to form a grammatically correct sentence. The scrambled sentences were simple four-word propositions without moral implications. Of main interest for data analyses were the response times for judgments of the second event in each pair in the experimental conditions, which will be referred to as "target" judgments.

RESULTS

Omission of Errors and Outliers

Before any statistical analyses were conducted, the data were examined for erroneous responses. Of the 504 target judgments in the experimental conditions, 19 (3.8%) were judged by participants as irreproachable, and were excluded from subsequent analyses. Fourteen of the primes were misclassified by participants (judged immoral when intended to be irreproachable, and vice versa), which rendered the exclusion of 2.8% of the prime events. Target response times were standardized within each category and experimental condition to identify potential outliers. In total, ten responses (2%) with standard deviations between 2.7 and 4.2 above the mean were excluded from analyses, meeting the recommendations of Bargh and Chartrand [33].

Facilitation Effects

The variables of main interest were the response times required for events when presented in the target conditions compared with response times for the same events when presented as primes (i.e., first in a pair). The size of the difference indicates the extent to which responses to target events were facilitated (or inhibited) as a function of the preceding primes. In order to test the hypothesized facilitation effects, planned contrasts were computed which compared each participant's mean response time for events in the prime position to the mean response times for events in each target condition. As predicted, responses to events in the SAME condition were significantly faster ($M = 2434$ ms, $SD = 513$) than responses to events in the prime position ($M = 2691$ ms, $SD = 684$), $F(1, 20) = 6.23$, $p < .05$, $\eta^2 = .24$. Also as predicted, the responses in the DIFFERENT condition ($M = 2521$ ms, $SD = 356$) were not significantly facilitated, $F(1, 20) = 2.32$, $p = .14$, $\eta^2 = .10$. Similarly, there was no significant facilitation effect for events in the FILLER condition ($M = 2580$ ms, $SD = 648$), $F(1, 20) = 2.37$, $p = .14$, $\eta^2 = .11$. The facilitation effects are presented in Fig. (3).

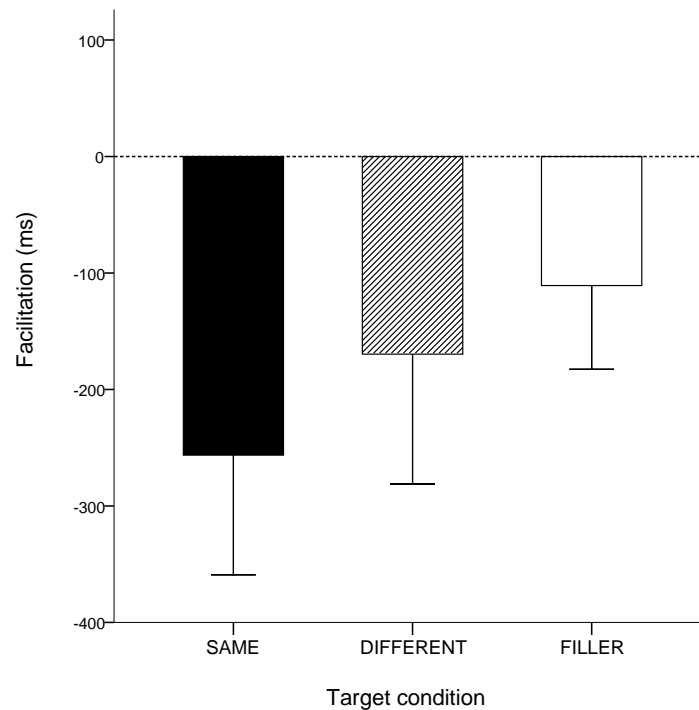


Fig. (3). Mean facilitation scores (+SE) as a function of target condition in Experiment 2.

Discussion

The results of Experiment 2 demonstrate that the exposure to immoral events speed up the processing of conceptually related events. This facilitation effect only occurred when the prime event was taken from the same category—as identified by Biel *et al.* [21]—as the target event. Since target responses were not facilitated by exposure to immoral events from a different category, the effect cannot be attributed to a priming of moral responding in general. This finding fits well with a large body of research showing that the strength of association between mental representations is a strong determinant of the magnitude of priming effects [32, 35].

The present findings also give support to Biel *et al.*'s [21] distinction between categories of immoral events. Representations of immoral events are assumed to be stored in memory as instances of violations of a specific moral value (see Fig. 2), and different values are associated with different types of events. Although the exact definitions of these values remain to be fully explored, our findings show that there is psychological validity in the proposed distinction. In sum, it appears that there is a semantic relatedness within, and distinctiveness between, classes of events with moral implications.

GENERAL DISCUSSION

The typicality effect demonstrated in Experiment 1 and the facilitation effect in Experiment 2 show that moral reactions display properties typical of automatic judgments [16]. First, the typicality of encountered instances is positively related to the speed of responding. Thus, in the same way as prototypical objects (e.g., robin—bird) are more readily

categorized than less prototypical ones (e.g., ostrich—bird), typical instances of immoral events are more readily classified as immoral than are atypical instances. This implies that immoral events are subject to a typicality gradient similar to other prototypically stored concepts [30, 31]. Second, as is the case with attitudes [e.g., 36] personality traits [e.g., 37], and norms [e.g., 38], representations of immoral events are susceptible to variations in accessibility. By temporarily increasing the accessibility of a certain category of immoral events, judgments with relevance to that category were sped up. Third, the strength of accessibility effects appears to be dependent on semantic relatedness [35]. Priming a specific category of immoral events facilitated responses to events violating the same moral value, but did not affect responses to immoral events in general. These findings cannot be accounted for by traditional rationalist models of moral judgment [e.g., 3] as they are not designed to explain immediate reactions to obvious moral transgressions. Neither of the above effects would be predicted from the process of conscious analysis and integration of situational information with moral principles and values that is proposed in earlier theories. To fully explain the present results, one must assume the existence of preexisting memory representations of immoral events. More recently the moral rules model of Darley & Shultz [39] acknowledges that moral judgments may be based on learned mental constructs. However, the cognitive architecture of these “rules” is not described on a level of specificity necessary to make meaningful comparisons with the present data.

A tentative framework for the cognitive architecture of moral value representations has been proposed by Biel *et al.* [21] (see Fig. 2). According to this view, intuitive moral reactions stem from a brief matching between the encoun-

tered situation and a corresponding mental prototype. In our view, however, every encountered immoral event is not likely to produce an automatic moral response. If the event is unfamiliar or ambiguous (i.e., not represented by a prototype) no match is expected, hence necessitating a more thorough, time-consuming cognitive analysis (see Fig. 1). This process is probably more akin to the process suggested in the traditional rationalist models. However, the present study concerned only reactions of an automatic nature.

An alternative interpretation of the data can be made concerning the level of specificity on which the matching process occurs. It is possible that encountered events are not matched against the specific attributes contained within a prototype, but rather at a more abstract schema level. This would necessitate an abstraction process whereby attributes are transformed into more general elements in order to allow for a comparison against a superordinate schema. Response time differences could then result from the fact that elements are more easily extracted from typical events than from atypical events. The design of Experiment 1 does not fully preclude such explanation. However, based on previous research demonstrating that mental representations of everyday events are generally stored as prototypes at a highly specific level [27, 28], the prototype-matching explanation seems more plausible.

The findings of the present study are consistent with recent developments in the area of moral judgment where dual-process theories have been incorporated into the field [19, 20, 40]. However, previous research on the relationships between moral reactions and information processing has been relatively vague as to the exact mechanisms underlying reactions at the automatic end of the automatic–controlled continuum. The present analysis provides a tentative outline for the cognitive basis of such reactions, and the aim of this research is to specify the constituents and workings of moral cognition at a more specific level than previously. The results of the present study demonstrate that adopting a prototype view of immoral event representation can further our understanding of the cognitive processes underlying moral reactions.

ACKNOWLEDGEMENTS

This research was partly funded by a grant from Lars Hierta Memorial Foundation awarded to the second author.

REFERENCES

- [1] Haan N. Can research on morality be "scientific"? *Am Psychol* 1982; 37: 1096-104.
- [2] Haan N, Aerts E, Cooper BA. *On moral grounds: The search for practical morality*. New York: New York University Press 1985.
- [3] Kohlberg L. Moral stages and moralization. In: Lickona T, Ed. *Moral development and behavior: theory, research, and social issues*. New York: Holt, Rinehart & Winston 1976; pp. 31-53.
- [4] Kohlberg L. A current statement on some theoretical issues. In: Modgil S, Modgil C, Eds. *Lawrence Kohlberg: Consensus and controversy*. Lewes: Falmer Press 1986; pp. 485-546.
- [5] Schwartz SH. Are there universal aspects in the structure and contents of human values? *J Soc Issues* 1994; 50: 19-45.
- [6] Feather NT. Values, valences, and choice: The influences of values on the perceived attractiveness and choice of alternatives. *J Pers Soc Psychol* 1995; 68: 1135-51.
- [7] Feather NT. Values, valences, expectations, and actions. *J Soc Issues* 1992; 48: 109-124.
- [8] Gilligan C. *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press 1982.
- [9] Kohlberg L. *The philosophy of moral development: Moral stages and the idea of justice*. San Francisco: Harper & Row 1981.
- [10] Piaget J. *The moral judgment of the child*. New York: Free Press 1965.
- [11] Rohan MJ. A rose by any name? The values construct. *Pers Soc Psychol Rev* 2000; 4: 255-77.
- [12] Schwartz SH, Bilsky W. Toward a universal psychological structure of human values. *J Pers Soc Psychol* 1987; 53: 550-62.
- [13] Schwartz SH, Bilsky W. Toward a theory of the universal content and structure of values: extensions and cross-cultural replications. *J Pers Soc Psychol* 1990; 58: 878-91.
- [14] Chaiken S, Trope Y, Eds. *Dual-process theories in social psychology*. New York: Guilford Press 1999.
- [15] Evans JStBT. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol* 2008; 59: 255-278.
- [16] Bargh JA. Principles of automaticity. In: Higgins ET, Kruglanski A, Eds. *Social psychology: Handbook of basic principles*. New York: Guilford 1996; pp. 169-83.
- [17] Bargh JA. The four horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In: Wyer RS, Srull TK, Eds. *Handbook of social cognition*, 2nd. ed. Hillsdale, NJ: Erlbaum 1994; pp. 1-40.
- [18] Greene JD, Morelli SA, Lowenberg K, Nystrom LE, Cohen JD. Cognitive load selectively interferes with utilitarian moral judgment. *Cognition* 2008; 107: 1144-54.
- [19] Haidt J, Koller SH, Dias MG. Affect, culture, and morality, or is it wrong to eat your dog? *J Pers Soc Psychol* 1993; 65: 613-28.
- [20] Sunstein CR. Moral heuristics. *Behav Brain Sci* 2005; 28: 531-73.
- [21] Biel A, Fransson N, Dahlstrand U. Moral values assigned to real-life events. *Scand J Psychol* 1997; 38: 21-8.
- [22] Medin DL. Concepts and conceptual structure. *Am Psychol* 1989; 44: 1469-81.
- [23] Murphy GL, Medin DL. The role of theories in conceptual coherence. *Psychol Rev* 1985; 92: 289-316.
- [24] Mandler G. Approaches to a psychology of value. In: Hechter M, Nadel N, Michod RE, Eds. *The origin of values*. New York: Walter de Gruyter 1993; pp. 229-58.
- [25] Fransson N, Biel A. Morality and norm violation. *Gothenburg, Sweden: University of Gothenburg, Department of Psychology. Göteborg Psychological Reports* 1997; Vol. 27: No 3.
- [26] Reed SK. Pattern recognition and categorization. *Cogn Psychol* 1972; 3: 382-407.
- [27] Laljee M, Lamb R, Abelson, RP. The role of event prototypes in categorization and explanation. In: Stroebe W, Ed. *European review of social psychology*, Oxford, UK: John Wiley & Sons 1992; Vol. 3. pp. 153-82.
- [28] Trafimow D, Wyer RS. Cognitive representation of mundane social events. *J Pers Soc Psychol* 1993; 64: 365-76.
- [29] Wyer RS, Srull TK. *Memory and cognition in its social context*. Hillsdale, NJ: Lawrence Erlbaum 1989.
- [30] Hampton JA. Similarity-based categorization and fuzziness of natural categories. *Cognition* 1998; 65: 137-65.
- [31] Rosch E. Cognitive representations of semantic categories. *J Exp Psychol Gen* 1975; 104: 192-232.
- [32] Baddeley A. *Human memory: theory and practice*. Hove, UK: Psychology Press 1997.
- [33] Bargh J A, Chartrand TL. The mind in the middle: A practical guide to priming and automaticity research. In: Reis HT, Judd CM, Eds. *Handbook of research methods in social and personality research*. Cambridge, UK: Cambridge University Press 2000; pp. 253-85.
- [34] Anderson JR. *Cognitive psychology and its implications*, 3rd ed. New York: Freeman 1990.
- [35] Higgins ET. Knowledge activation, accessibility, applicability, and salience. In: Higgins ET, Kruglanski AW, Eds. *Social psychology: Handbook of basic principles*. New York: Guilford Press 1996; pp. 133-68.
- [36] Fazio RH, Sanbonmatsu DM, Powell MC, Kardes FR. On the automatic activation of attitudes. *J Pers Soc Psychol* 1986; 50: 229-38.

- [37] Higgins ET, Rholes WS, Jones CR. Category accessibility and impression formation. *J Exp Soc Psychol* 1977; 13: 141-54.
- [38] Kallgren CA, Reno RR, Cialdini RB. A focus theory of normative conduct: When norms do and do not affect behavior. *Pers Soc Psychol Bull* 2000; 26: 1002-12.
- [39] Darley JM, Shultz TR. Moral rules: Their content and acquisition. *Annu Rev Psychol* 1990; 41: 525-56.
- [40] Haidt J. The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychol Rev* 2001; 108: 814-34.

Received: Februar 23, 2010

Revised: April 19, 2010

Accepted: April 20, 2010

© Fransson and Ask; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.