Competent Deontic Reasoning: The Abstract Deontic Selection Task Revisited

Sieghard Beller (beller@psychologie.uni-freiburg.de)
Andrea Bender (bender@psychologie.uni-freiburg.de)
Center for Interdisciplinary Research (ZiF), Bielefeld University
D-33615 Bielefeld, Germany
Department of Psychology, University of Freiburg
D-79085 Freiburg, Germany

Abstract

The abstract deontic selection task was developed with the aim of demonstrating abstract rule use in a specific domain (i.e., deontic rules). Yet, the solution rate, while being substantially higher than with abstract *non-deontic* tasks, did not reach the level obtained with *concrete* deontic tasks. What are the reasons for this—difficulties with abstract rule use? A task analysis based on deontic principles uncovers several problems with the formulation of the task. Three experiments replicate the difficulties with the original task and show that performance increases, when the problems are resolved. The results provide novel insights into the interpretation of deontic rules and into the role that such content-specific, but abstract tasks can play for the study of reasoning processes.

Keywords: Deontic reasoning; social rules; deontic selection task; dual process theory; pragmatic reasoning schemas.

Introduction

Abstract reasoning is typically considered as more difficult and error prone than contextualized reasoning. A widely accepted and often cited case in point is *Wason's Selection Task:* The classical abstract task demands testing whether an arbitrary conditional statement is true or false (Wason, 1966) and is solved correctly only by few participants (10-20%). In contrast, structurally similar but contextualized versions that demand testing whether or not a concrete deontic conditional is being followed (e.g., Cox & Griggs, 1982) are solved correctly by the majority of participants (about 75%; Evans, 2003, p. 456; cf. Beller, 2010).

According to the *dual-process account* (Evans, 2008), abstract reasoning is characterized as domain-general, rule-based, sequential, controlled, and slow, and is limited by working memory capacity ("System 2 reasoning"); contextualized reasoning, on the other hand, is characterized as domain-specific, associative, parallel, automatic, and rapid, and is independent of working memory ("System 1 reasoning"). However, this distinction is often not clear cut: On the one hand, there are many examples for people making use of heuristic, associative System 1 processes when thinking about abstract, de-contextualized tasks; in the classical abstract selection task, for instance, they tend to apply the matching heuristic (Evans, 2003). On the other hand, people can also engage in analytic, rule-based System 2 processes when thinking about concrete, contextualized tasks as argued in Beller and Spada (2003).

The abstract deontic selection task is an interesting hybrid: It is domain-specific and contextualized, as it refers to a domain and a contextual framing we all are familiar with (to search for violators of a deontic regulation). But, instead of presenting a concrete, familiar rule like the *Drinking Age* rule "If a person is drinking beer, then that person must be over 16 years of age" (Cox & Griggs, 1982), which may activate an instance-based mode of reasoning, the task is formulated in an abstract way. Finding people's reasoning performance to be as accurate as with concrete, familiar material would thus provide evidence for abstract rule use in a specific domain (Smith, Langston & Nisbett, 1992).

The original version, introduced as "permission problem", reads as follows (Cheng & Holyoak, 1985, p. 403):

> *Suppose you are an authority checking whether or not people are obeying certain regulations. The regulations all have the general form, "If one is to take action A, then one must first satisfy precondition P." In other words, in order to be permitted to do "A", one must first have fulfilled prerequisite "P". The cards below contain information on four people: One side of the card indicates whether or not a person has taken action "A", the other indicates whether or not the same individual has fulfilled precondition "P". In order to check that a certain regulation is being followed, which of the cards below would you turn over? Turn over only those that you need to check to be sure.*
>
> *Card (1) "Has taken action A"*
> *Card (2) "Has not taken action A"*
> *Card (3) "Has fulfilled precondition P"*
> *Card (4) "Has not fulfilled precondition P"*

Which cards does one need to check? This question is not too difficult to answer: The regulation is being followed if a person, who takes action "A", has fulfilled precondition "P". Therefore two cards need to be checked: Card (1) "Has taken action A" in order to find out whether this person has fulfilled precondition "P", and Card (4) "Has not fulfilled precondition P" in order to exclude that this person has taken action "A", as otherwise the regulation would be violated.

The task is about an abstract regulation. This should pose no problems as, according to *Pragmatic Reasoning Schema Theory* (Cheng & Holyoak, 1985), people possess abstract schemas for handling such rules. People should thus be able to solve the abstract deontic task as easily as they solve concrete ones. However, this is not the case as indicated by the meta-analysis reported in Beller (2010, p. 127): Instead, the

solution rate drops from 73.8% in concrete deontic selection tasks (N = 1,010; 26 Experiments) to 58.4% in abstract deontic tasks (N = 320; 10 Experiments), indicating—at first glance—some difficulties with abstract rule use in the sense of Smith, Langston and Nisbett (1992).

According to *Social Contract Theory* (Cosmides, 1989), this result comes not as a surprise because the abstract deontic task lacks a clear cost-benefit structure necessary to identify rule violators: persons who take the benefit (e.g., the beer) without "paying the costs" (i.e., fulfilling the age requirement). But if cost-benefit information were indeed necessary for the solution, should we then not expect a much stronger decrease in the solution rate than the observed 15%? And if, on the other hand, missing cost-benefit information is not the reason, what else then could be responsible for the reduced solution rate? In order to answer these questions, the "permission problem" needs to be analyzed in more detail.

A Task Analysis

How to Solve the Task

As suggested by Beller (2008), the first step in solving a deontic task is to identify which action constraint is imposed by a deontic rule (*constraint principle*). In the "permission problem", the restriction concerns action "A", and the condition is "P". According to the description, condition "P" is necessary for the permission to do "A". Consequently, if "P" is not fulfilled then action "A" is forbidden. Assuming that "P" is the only relevant condition (*exhaustivity*), the deontic constraint can then be represented according to the *equivalence principle* by the ban "If condition P is not fulfilled then action A is forbidden; otherwise it is allowed":

Ban (B): $\neg$condition_P $\leftrightarrow$ forbidden(action_A)

After the deontic action constraint has been identified, the second step consists of drawing the appropriate deontic inferences. From a deontic norm like ban (B), three types of inferences are possible (Beller, 2008): *Forward inferences* from the condition side of the norm to the deontic status of the regulated action (i.e., whether it is forbidden or allowed, obligatory or not), *backward deontic-to-factual inferences* from the deontic status of the regulated action to the condition (i.e., whether or not the condition is fulfilled), and *backward factual-to-deontic inferences* from information on whether or not the regulated action is taken to the deontic status of the condition (e.g., whether or not it is deontically necessary).

In order to come up with the correct solution to the selection task ("A & not-P"), one has to check *each* of the persons and to infer whether he or she might have *violated* the rule (i.e., has performed the banned action "A" without fulfilling condition "P"). The inference process might proceed as follows: For card (1) "Has taken action A", one can conclude by a backward factual-to-deontic inference that this person must fulfil condition "P"; otherwise the rule is violated. For card (2) "Has not taken action A", one can conclude by corresponding backward inferences that this person need not, but may fulfil condition "P". Therefore, the rule cannot be violated. For card (3) "Has fulfilled condition P", one can conclude by forward inferences that this person may, but need not take action "A"; again, the rule cannot be violated. Finally, for card (4) "Has not fulfilled condition P", one can infer by a corresponding forward inference that this person may not take action "A"; otherwise the rule is violated.

Problems in the Original Task

Unfortunately, Cheng and Holyoak's (1985) original task is formulated in a way that renders it difficult to interpret the deontic regulation and to understand the task requirement correctly, with direct effects on how the task is solved.

With regard to the interpretation of the deontic regulation, several formulations are problematic: First, by stating that "The regulations *all* have the general form ..." (italics added), the "permission problem" suggests that not a single rule has to be checked, but a set of *possibly different* rules. Second, the formulation "If one is to take action A ..." can be understood as referring to an intention ("if someone *wants* to take action A ..."). However, it is not the *intention* that is deontically constrained, but *doing* the action. Third, the condition is described as a composition of "P" and an additional temporal constraint: "P" has to be fulfilled *prior* to action "A". This might initiate some temporal reasoning, which cannot be resolved unequivocally, as the necessary temporal information is missing on the cards. Fourth, the precondition has to be fulfilled prior to the action, but the regulation mentions the two elements in the reverse order. And finally, the regulation qualifies condition "P" as *necessary* with the modal *must*, but does not specify whether "P" is *sufficient* for the permission to take action "A".

With regard to the task requirement, two formulations are problematic: The first is the instruction "to check that a *certain* regulation is being *followed*" (italics added). It emphasizes that one particular regulation has to be checked. At the very beginning, the task speaks of regulations in the plural form, which may cause uncertainty about which specific regulation out of this set is meant. More severely, this instruction emphasizes rule *following*, whereas concrete selection tasks like the *Drinking Age Problem* focus on rule *violation*. This may appear to be a subtle distinction, but in fact it is one that makes an important difference: Whether a person does follow the regulation can be checked by turning over the "A"-card (1), but not by turning over the "not-P"-card (4), as rule following is relevant only for persons to which the rule *applies*: Person (1), who has taken the critical action "A", is subject to the rule. If it turns out that this person has fulfilled prerequisite "P", it is clear that he or she has *followed* the rule; otherwise he or she has *violated* it. For person (4), who has not fulfilled prerequisite "P", the case is different: If this person has taken action "A", it is clear that this person is subject to the rule and has *violated* it, but if this person has *not* taken action "A", the rule is simply *not applicable*, and hence it is not possible to detect rule following. People, therefore, may neglect the "not-P"-card and select only the "A"-card. The second problematic formulation is the instruction "Turn over only those [cards] that you need to

check to be sure". This might cause people to be *cautious* not to select too many cards. Some might think that finding one rule follower (or rule violator) would be sufficient.

In summary, the formulation of the task makes it difficult to extract the relevant deontic information, and the instruction induces a general preference for choosing only few cards (caution) as well as a specific preference for the "A"-card alone (to check rule following). In combination, this leads to a reduction of selection task performance and, accordingly, to an underestimation of people's deontic competencies. By eliminating the problematic formulations, performance should increase. In the following, three experiments test this hypothesis. The results will be discussed with regard to the role this kind of content-specific, but abstract tasks can play in gaining new insights into content-specific reasoning.

Three Experiments

In each of the experiments (taken from Beller, in press, and described here together for reasons of space), the original task from Cheng and Holyoak is compared to four new ones. By varying the formulation of the deontic rule (strong vs. weak; backward vs. forward), the question of *rule interpretation* is addressed; the problems of the *instruction* are addressed by focusing on rule violation instead of rule following.

The new tasks were constructed according to the schema shown below. Clues about intentions and temporal conditions are avoided; the instruction emphasizes that each person should be checked for rule violation; and explicit negatives are used to clarify when action "A" is *not* taken and when condition "P" is *not* fulfilled:

> *Imagine you are a member of an authority that checks whether people conform to or violate a particular rule. The rule is: {one of the new rules from below}. The "cards" presented below represent four persons. On one side of each card is written whether or not the respective person takes action "A", on the other side is written whether or not he or she fulfills condition "P". Your task: Indicate all cards that you have to turn over—all of which you need to know the information on the back—in order to find out whether the respective person violates the rule.*
>
> *Person (1) "Takes action A"*
> *Person (2) "Does not take action A"*
> *Person (3) "Fulfills condition P"*
> *Person (4) "Does not fulfill condition P"*

Each task refers to a single deontic rule, which is given without further clarifications. In Experiment 1, the following four rules were used:

> *Obligation O1:* "If a person takes action A, then he or she *must* fulfill condition P."
>
> *Release R1:* "If a person does not take action A, then he or she *need not* fulfill condition P."
>
> *Ban B1:* "If a person does not fulfill condition P, then he or she *must not* take action A."
>
> *Permission P1:* "If a person fulfills condition P, then he or she *may* take action A."

All four rules can logically be derived from ban (B), with obligation O1 being analogous to the original rule. Two rules use a strong deontic modal and describe a deontic constraint explicitly (O1 and B1), the two other rules use a weak deontic modal (R1 and P1). Taken literally, these latter rules cannot be violated in the deontic sense, as they do not express a deontic constraint explicitly.

For Experiment 2, these rules were extended by one sentence that make their complementary side explicit in order to facilitate to derive ban (B) from the weak rules:

> *Obligation O2:* "If a person takes action A, then he or she *must* fulfill condition P; otherwise he or she need not fulfill it."
>
> *Release R2:* "If a person does not take action A, then he or she *need not* fulfill condition P; otherwise he or she must fulfill it."
>
> *Ban B2:* "If a person does not fulfill condition P, then he or she *must not* take action A; otherwise he or she may take it."
>
> *Permission P2:* "If a person fulfills condition P, then he or she *may* take action A; otherwise he or she must not take it."

In Experiment 3, it was manipulated, how the deontic modalities are expressed in the rules: not by deontic modals (e.g., *must not* or *may*) as in Experiment 1 and 2, but by semantically equivalent deontic verbs (e.g. to *forbid* or *permit*). With the verbs, explicit negations can be avoided, thereby eliminating another potential difficulty:

> *Obligation O3:* "If a person takes action A, then he or she is *obliged* to fulfill condition P; otherwise he or she is released from fulfilling condition P."
>
> *Release R3:* "If a person does not take action A, then he or she is *released* from fulfilling condition P; otherwise he or she is obliged to fulfill condition P."
>
> *Ban B3:* "If a person does not fulfill condition P, then action A is *forbidden;* otherwise action A is permitted."
>
> *Permission P3:* "If a person fulfills condition P, then action A is *permitted;* otherwise action A is forbidden."

For the original task, it is expected that people tend to avoid selecting too many cards (*caution hypothesis*) and tend to neglect the "not-P" card and to select the "A"-card alone (*rule-following hypothesis*) due to the various problems in the formulation of this task. For the new tasks, it is expected that people infer ban (B) more easily due to the clearer formulation. With ban (B) in mind, they should select the cards "A" and "not-P" when asked to check individuals for rule violation, independent of how the rule was formulated (*rule-violation hypothesis*). If asked for the literal meaning of the rules, however, people should differentiate between strong rules that can be violated by a person and weak rules that cannot be violated (*rule-evaluation hypothesis*).

Method

Materials. In each experiment, five deontic selection tasks were used: the original task and four new tasks. The new tasks were constructed according to the general schema shown above and differed only in the way the rule was formulated. In addition to the selection tasks, each experiment was supplemented with a second task. In Experiment 1, a rule *evaluation task* was used in order to check which of the rules O1, R1, B1, and P1 participants consider as violable in a strict deontic sense. The instruction asked participants to *consider for each statement whether it expresses a rule that can in fact be violated by a person.* In Experiment 2, this instruction was used to compare the weak rules from Experiment 1 with the corresponding explicit rules from Experiment 2 (i.e. R1, P1, R2, and P2). In Experiment 3, participants were asked to rank the rules O3, R3, B3, and P3 with respect to comprehensibility (from 1 = best to 4 = worst).

Participants. All participants were from the University of Freiburg and had no prior experience with the selection tasks. In Experiment 1, a total of 195 students (39 per condition) volunteered to participate for research credit (72 male, 123 female; mean age 23.2 years, $SD = 3.57$). In Experiment 2, a total of 125 students (25 per condition) volunteered to participate for research credit (36 male, 89 female; mean age 23.1 years, $SD = 5.27$). In Experiment 3, a total of 130 students (26 per condition) volunteered to participate as part of a study on a different subject for which they were paid (37 male, 91 female; mean age 23.5 years, $SD = 3.94$).

Design and Procedure. These were the same for all three experiments: Each participant received a booklet with general instructions, one selection task, and the additional task (on separate pages). In both tasks, two orders of answer options were permuted equally frequently across conditions. The different versions of the booklet (with the different selection task versions varying between subjects) were randomly assigned to participants. They were investigated in small groups, were instructed to work on the tasks in the given order, and were granted as much time as they needed.

Results

In the original selection task, people were expected to be cautious not to select too many cards, which should decrease the mean number of selected cards. With the focus on rule following, people should tend to neglect the "not-P" card and to select the "A"-card alone instead of the cards "A & not-P". In the new tasks, performance should improve with respect to all these indicators; and exactly this was found.

Overall analyses of the selection task data. In Experiment 1, the five experimental groups differed significantly in the mean number of chosen cards ($F(4, 190) = 3.99$; $p = .004$) as indicated by an analysis of variance, and also in the overall frequency of the "not-P"-card ($L^2 = 34.4$; $N = 195$; $df = 4$; $p < .001$), in the frequency of the "A"-card alone ($L^2 = 14.2$; $N = 195$; $df = 4$; $p = .007$), and in the frequency of the combination "A & not-P" ($L^2 = 36.2$; $N = 195$; $df = 4$; $p < .001$) as indicated by the likelihood ratios. Almost all effects were in the predicted directions (cf. Figure 1).

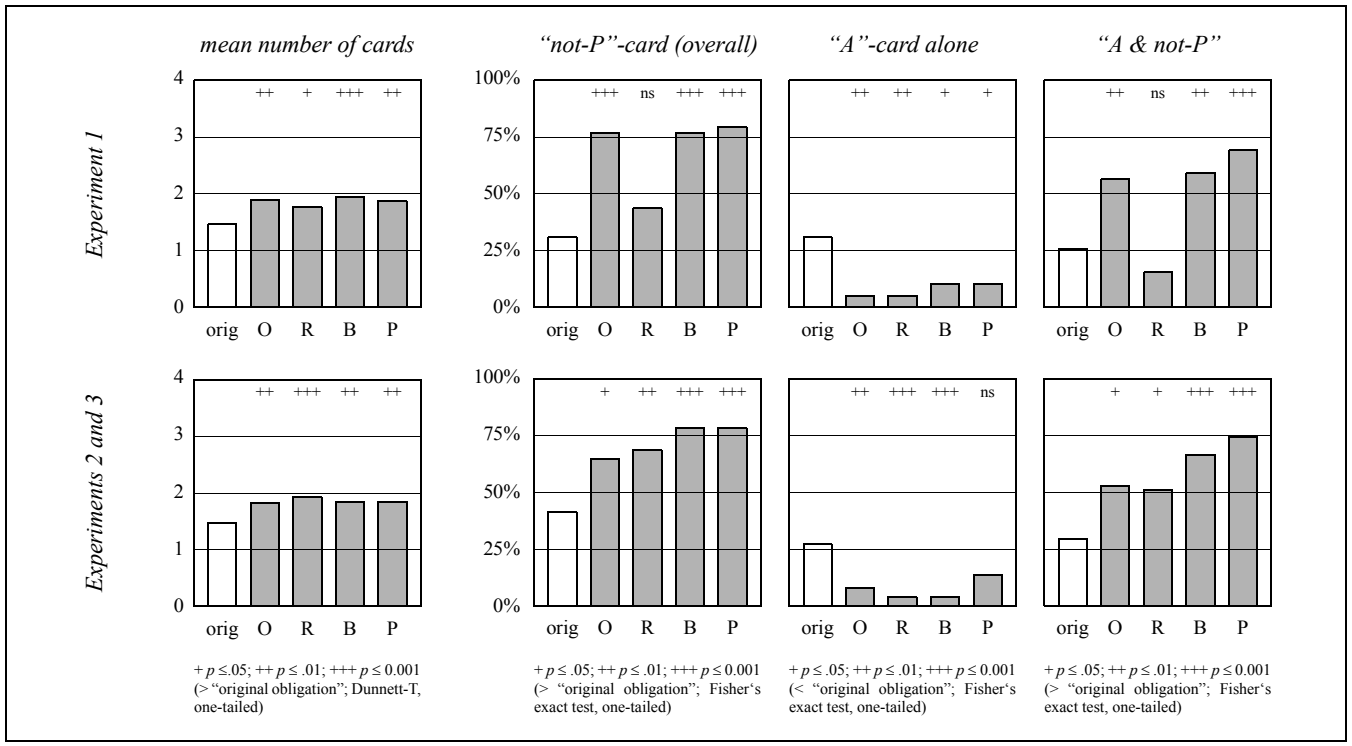

Figure 1: Four deontic indices for the original selection task (white bars) and the four new tasks Obligation, Release, Ban, and Permission (grey bars) in Experiment 1 (top row; $N = 195$) and aggregated across Experiments 2 and 3 (bottom row; $N = 255$).

As explained above, the selection tasks used in Experiment 2 and 3 differed only in one aspect: in how the deontic modality was formulated. They were thus analyzed jointly by means of an analysis of variance and log-linear analyses (Kennedy, 1992). As none of these analyses indicated a main effect "Experiment" nor an interaction "Experiment" × "Condition" (for all *Fs:* $p > .544$; for all L^2: $p > .553$), aggregating the data over the two experiments seemed justified. This also means that it made no difference for the solution of the tasks whether the deontic modality in the rule was expressed by modals with explicit negation (Experiment 2) or by verbs without explicit negation (Experiment 3). As in Experiment 1, the five experimental groups differed in the mean number of chosen cards ($F(4, 245) = 4.75$; $p = .001$), in the overall frequency of the "not-P"-card ($L^2 = 20.9$; $N = 255$; $df = 4$; $p < .001$), in the frequency of the "A"-card alone ($L^2 = 18.1$; $N = 255$; $df = 4$; $p = .001$), and in the frequency of the combination "A & not-P" ($L^2 = 25.2$; $N = 255$; $df = 4$; $p < .001$). Again, almost all effects were in the predicted directions (cf. Figure 1).

Original Selection Task. As predicted by the *caution hypothesis*, participants selected too few cards, that is, less than 2.0, the value expected according to the correct "A & not-P" solution (Experiment 1: 1.462; $t(38) = –6.06$; $p < .001$; Experiments 2&3: 1.471; $t(50) = –5.17$; $p < .001$). As predicted by the *rule-following hypothesis*, the main problem was to recognize the relevance of the "not-P"-card (Experiment 1: 30.8% "not-P"; Experiments 2&3: 41.2%). The proportion of the "A"-card alone was rather high (Experiment 1: 30.8%; Experiments 2&3: 27.5%), and the frequency of the correct deontic solution "A & not-P" was low (Experiment 1: 25.6%; Experiments 2&3: 29.4%).

New Selection Tasks. The performance in the new tasks differed significantly from the original task in all aspects: The mean number of chosen cards was higher in all four conditions (Experiment 1: 1.872 on average; Experiments 2&3: 1.858). The relevance of the "not-P"-card was clearly recognized (Experiment 1: in three out of four conditions; Experiments 2&3: in all four conditions), the proportion of the "A"-card alone was reduced (Experiment 1: in all four conditions; Experiments 2&3: in three out of four conditions) and, complementarily, the correct combination "A & not-P" increased (Experiment 1: in three conditions; Experiments 2&3: in all four conditions). This pattern indicates that, in most conditions, the majority of participants inferred ban (B) and identified rule violators more often than in the original task, as predicted by the *rule-violation hypothesis*.

Additional Tasks. The additional task in Experiment 1 compared the weak rules permission P1 and release R1 with the two strong rules ban B1 and obligation O1. According to the *rule-evaluation hypothesis*, participants should indicate that weak rules cannot be violated in the deontic sense, only strong rules can. To test this hypothesis, a strength-index was calculated by adding 1 for each strong rule that a person marked as violable, subtracting 1 for each weak rule, and dividing the result by 2. This renders a maximum score of 1 if the two strong rules were marked only, a minimum score of –1 if the two weak rules were marked only, and a score of 0 if strong and weak rules were marked in a balanced way. An analysis of variance indicated that the strength-indices of the five selection task conditions did not differ from one another ($F(4, 190) = .491$; $p = .742$). The overall index was positive (.667) and different from zero ($t(194) = 22.3$; $p < .001$). The data clearly support the *rule-evaluation hypothesis:* The strong rules were regarded as violable by most participants (obligation: 80.0%; ban: 84.6%), the weak rules were not (release: 5.6%; permission: 25.6%; $N = 195$).

The additional task in Experiment 2 compared the two weak rules from Experiment 1 (P1 and R1) that left the underlying deontic constraint implicit with the two explicit rules from Experiment 2 (P2 and R2) that expressed this constraint. A strength-index was calculated analogously to Experiment 1. If participants considered the explicit rules as the only violable ones then a high positive value should result. An analysis of variance indicated that the strength-indices of the five selection task conditions did not differ from one another ($F(4, 120) = .612$; $p = .655$). The overall index was positive (.652) and different from zero ($t(124) = 14.9$; $p < .001$). The data again support the *rule-evaluation hypothesis*: The explicit rules were regarded as violable by most participants (R2: 79.2%; P2: 81.6%), the implicit rules were not (R1: 11.2%; P1: 19.2%).

The additional task in Experiment 3 asked participants to rank the four rules from "easiest to understand" (= 1) to "most difficult to understand" (= 4). Two rules expressed an obligation and were formulated as backward rules: Obligation O3 focused on the obligation in the if-then clause, release R3 in the otherwise clause. The two other rules expressed a ban and were formulated as forward rules: Ban B3 focused on the ban in the if-then clause, permission P3 in the otherwise clause. All rules thus *expressed* a deontic constraint. But were they therefore equally easy to understand, too? In order to answer this question, an analysis of variance on the rank values was conducted with one within-subject factor "rule" (rank value assigned to each of the rules) and one between-subjects factor "condition" (the five selection task conditions). As comprehension of the rules should not depend on the type of selection task that individuals had solved before, an effect "condition" was not expected—and was not found either ($F(4, 124) = .990$; $p = .416$). But, a strong rule effect was found ($F(3, 372) = 210.6$; $p < .001$) with a clear rank order: Permission was regarded as most comprehensible ($m = 1.26$), followed by the ban ($m = 2.19$), the obligation ($m = 2.92$), and the release from obligation ($m = 3.60$). An interaction ($F(12, 372) = 4.13$; $p < .001$) indicated a slight difference between the experimental groups: The one exception from the general order *permission* < *ban* < *obligation* < *release* occurred in the group that had solved the selection tasks with obligation O3; here, the obligation switched to the second position: *permission* < *obligation* < *ban* < *release*.

Conclusions

Any sound test of a psychological theory requires that we understand the applied behavioral paradigms in all relevant aspects. The task analysis presented in this paper (cf. Beller,

in press) indicated that a great deal of the difficulties with the original abstract deontic selection task arose from an inadequate formulation of the task. The experimental data confirm this analysis: They provide evidence for the specific difficulties with the original task and show that performance can be substantially improved, if the selection task is formulated more clearly with regard to its deontic nature.

The abstract deontic selection task was introduced in order to demonstrate that people possess abstract reasoning schemas for reasoning from deontic rules (Cheng & Holyoak, 1985). The reduced performance with the original version (as compared to concrete versions) was attributed by some scholars to lacking cost-benefit information, which is characteristic for social contract rules (e.g., Cosmides, 1989). This explanation can now be rejected: As all the tasks used in Experiments 1 to 3 have the same cost-benefit structure, the differences between the original task and the new tasks unequivocally show that the reduced performance in the original task must have other reasons.

What do the experimental data tell us about people's deontic competencies? Performance was best with tasks that combined a genuine violation instruction with a genuine permission rule: With 72.2% of participants choosing "A & not-P" on average across the three experiments, these tasks reliably approached the average result from the concrete deontic tasks (73.8%, according to Beller, 2010, p. 127). This finding may count as an indication for rule guided behavior in a specific content domain (Smith, Langston & Nisbett, 1992), that is, for System 2 reasoning in a domain-specific, contextualized task. Permission rules were also rated as most comprehensible even though rules like the one used in Experiment 1 do not express a deontic constraint explicitly. Consequently, before people can come up with the "A & not-P" solution in the selection task, they must infer from the rule which action is banned under which condition, and the data are consistent with the assumption that they do this according to the principles suggested in Beller (2008). As a consequence, the exact wording of the rule—whether it is formulated forwards or backwards, as permission or as obligation—is not relevant for the identification of rule violations, as long as people are able to infer the adequate deontic regulation. In this sense, the results constitute an instance of the rule-change phenomenon in showing that people infer cases of rule violation independently of their relation to the "logical form" of the conditional statement (Cosmides, 1989; cf. Beller, 2001; 2010; Beller & Spada, 2003).

Beyond that, tasks like the abstract deontic selection task are important tools for studying domain-specific reasoning processes as they combine two features: They are *content-specific* as they refer to a particular content domain (like deontic norms), and *abstract* as they do not refer to a specific instance, thereby avoiding content effects due to the experience that people have with particular instances (for a causal example see Beller & Kuhnmünch, 2007). As with content effects in general (cf. Beller & Spada, 2003), specific instances of deontic rules may facilitate performance, if experiences with a rule are available directly from memory (for an example see Beller, 2001). Likewise, experience with a specific instance might also suppress performance, if inhibiting aspects like additional conditions or additional social norms are evoked by its content (for an example see Beller, Bender & Song, 2009). Our theories of reasoning—even the content-specific ones—are not formulated for single instances (like a specific drinking-age rule), but refer abstractly to classes of comparable situations (like deontic norms in general). Tasks that are both, content-specific and abstract, are thus a mean for testing domain-specific theories in a purer way, and can provide a baseline that help us to detect and to assess the extent of such instance-specific content effects.

Acknowledgements. We are grateful to Lukas Bischof, Miriam Hansen, Gregory Kuhnmünch, and Nikol Rummel for helping with data collection, and to Lisa Hüther for valuable comments on earlier versions of this article.

References

Beller, S. (2001). A model theory of deontic reasoning about social norms. In J. D. Moore, & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum.

Beller, S. (2008). Deontic norms, deontic reasoning, and deontic conditionals. *Thinking and Reasoning, 14,* 305-341.

Beller, S. (2010). Deontic reasoning reviewed: psychological questions, empirical findings, and current theories. *Cognitive Processing, 11,* 123-132.

Beller, S. (in press). Concrete problems in the abstract deontic selection task—and how to solve them. *The Quarterly Journal of Experimental Psychology.*

Beller, S., Bender, A., & Song, J. (2009). Conditional promises and threats in Germany, China, and Tonga: Cognition and Emotion. *Journal of Cognition and Culture, 9,* 115-139.

Beller, S., & Kuhnmünch, G. (2007). What causal conditional reasoning tells us about people's understanding of causality. *Thinking and Reasoning, 13,* 426-460.

Beller, S., & Spada, H. (2003). The logic of content effects in propositional reasoning: The case of conditional reasoning with a point of view. *Thinking and Reasoning, 9,* 335-378.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391-416.

Cosmides L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-276.

Cox, R. J., & Griggs, R. A. (1982). The effect of experience on performance in Wason's selection task. *Memory & Cognition, 10,* 496-502.

Evans, J. St. B. T. (2003). In two minds: dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59,* 255-278.

Kennedy, J. J. (1992). *Analyzing qualitative data*. New York: Praeger.

Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case for rules in reasoning. *Cognitive Science, 16,* 1-40.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology.* Harmondsworth, UK: Penguin.