Propositional reasoning and working memory: the role of prior training and pragmatic content
1. Introduction

The concept of working memory has been used in explaining deductive reasoning performance from different theoretical viewpoints. Nevertheless, only few studies have directly investigated the dependence of reasoning performance on working memory resources (see Gilhooly, 1998, for an overview). The experiments reported herein focused on the involvement of working memory in propositional reasoning and on the moderating effects of training and task content. By employing different kinds of training, the experiments aimed at probing the cognitive processes underlying propositional reasoning according to three classes of theories, that is, mental logic theories, the mental model theory, and context-specific theories of reasoning. Following recent dual process models of human thinking (e.g., Evans & Over, 1996; Sloman, 1996; Stanovich, 1999), the training conditions were expected to reduce the impact of non-analytic, heuristic processes and to strengthen the use of analytic, controlled deduction processes in task performance. Thus, the training conditions allowed us to investigate working memory demands of analytic reasoning processes, disentangled from the overpowering function of non-analytic processes and shortcuts.

1.1. Theories of propositional reasoning

The term “propositional reasoning” refers to conclusions that are drawn from premises with logical connectives like “and”, “or”, and “if ... then ...” which link elementary propositions. The formal definition of premises in propositional calculus can be represented at least in two different ways: as a set of inference rules or by means of truth tables. Inference rules describe deductions which result from, or lead to, a premise with a given connective. Truth tables indicate the truth value of a premise for each combination of true or false elementary propositions. The two ways of representing premises in formal logic are related to two classes of psychological theories of propositional reasoning, namely to mental logic theories and to the theory of mental models. A third class of psychological theories refer to context-dependent inference schemata evoked by particular content domains.

Mental logic theories postulate a universal repertory of natural inference rules that are used to derive or to prove conclusions from given premises. To apply the abstract inference rules to premises framed in a semantic context, a comprehension mechanism is required which maps semantic information onto syntactic schemata. Moreover, reasoning strategies are necessary that guide the application of inference rules, for instance by linking single rules to deduction chains, setting up subgoals, storing newly derived premises or assertions, and feeding them into the deduction process. Recent theories of propositional reasoning that specify a repertory of natural inference rules, comprehension mechanisms, and reasoning strategies have been proposed by Braine and colleagues (Braine & O’Brien, 1991, 1998; Braine et al., 1995; Braine, Reiser, & Rumain, 1984; Braine & Rumain, 1983) and by Rips (1983, 1994, 1995). These theories predict that a particular sequence of inference
steps is performed to reach a conclusion from a given set of premises, and that the number of inference steps, their accessibility, and the complexity of the reasoning strategies determine the difficulty of a reasoning problem (e.g., Braine et al., 1984; Rips, 1994).

In contrast to mental logic accounts, the theory of propositional reasoning by mental models (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991; Johnson-Laird, Byrne, & Schaeken, 1992; see also Evans, 1993) is based on semantic, rather than syntactic processes of deduction. According to the mental model theory, drawing a conclusion from a set of premises involves three steps. In the first step, an initial mental model is formed which represents a state of events that is compatible with the premises. In the second step, a putative conclusion is derived which is true in the initial mental model and which is informative, that is, does not state the obvious. In the third step, further mental models of the premises are sought to test for counterexamples to the putative conclusion. If no counter-example is found, the conclusion is finally accepted as valid. Otherwise the conclusion is rejected, and either the response “nothing follows” is given, or another putative conclusion is derived from the models considered so far. Although the set of potential mental models of a premise mirrors the cells of the truth table with value “true”, it is assumed that reasoners do not use complete truth tables, but that they rely on more parsimonious model representations, which are suited for reasoners’ limited mental capacities. Consequentially, the theory predicts that task difficulty is a function of the necessity to flesh out mental models beyond the initial model and of the number of mental models that must explicitly be taken into account to reach a valid conclusion (Johnson-Laird et al., 1992; Klauer & Oberauer, 1995).

The theories of mental logic and of mental models build on the assumption that the computational routines governing deductive reasoning are very general in nature. That is, deduction by inference rules and mental models, respectively, come into force irrespective of the contents of the premises. In contrast, a third class of theories postulate specific reasoning algorithms for certain content domains or situations. For instance, the theory of pragmatic reasoning schemata (Cheng & Holyoak, 1985, 1989; Cheng, Holyoak, Nisbett, & Oliver, 1986; Holyoak & Cheng, 1995) states that reasoning relies on generalized knowledge structures which are acquired by induction in situations reflecting permission, obligation, or causality relations. The knowledge structures, or schemata, contain context-sensitive production rules which, once established, determine the inferences in situations that trigger the schema. These rules are pragmatic, rather than syntactic, in that they serve the achievement of certain goals in a given context. The theory predicts facilitating effects of schema-related task contents (Cheng & Holyoak, 1985, 1989; Cheng et al., 1986) and social perspective (Holyoak & Cheng, 1995; Politzer & Nguyen-Xuan, 1992). Alternative theoretical approaches postulating domain-specific reasoning processes comprise the social contract theory (Cosmides, 1989; Cosmides & Tooby, 1992) and the hypothesis of a cheater detection algorithm (Gigerenzer & Hug, 1992), both of which specify inference algorithms for narrowly defined classes of social situations that supposedly were crucial in the evolutionary adaptation of mankind.
1.2. Propositional reasoning and working memory

The above theories differ with respect to the role of working memory in propositional reasoning. According to mental logic theories, reasoning requires working memory resources for the matching rules, which interpret verbal premises with regard to their syntactic structure, for the inference rules in the actual deduction process, and for storage operations on the premises and newly derived assertions (cf. Rips, 1995). Following the theory of reasoning by mental models, working memory limitations constrain the ability to represent larger sets of fleshed-out models, thus restricting the search for counter-examples to a putative conclusion (Johnson-Laird & Byrne, 1991; Johnson-Laird et al., 1992). Hence, both mental logic and mental model accounts contend that the cognitive processes of propositional reasoning are vulnerable to deficiencies of working memory resources. The theory of pragmatic reasoning schemata, in contrast, predicts that the evocation of an appropriate reasoning schema releases from cognitively demanding deductions (Cheng & Holyoak, 1985), and that reasoning performance depends on the fit between the premises and prevailing schemata yielding logically correct inferences, rather than on memory capacity (Cheng et al., 1986). Accordingly, reasoning on the basis of crystallized production rules, which are part of a pragmatic knowledge structure, should be less dependent on the availability of working memory resources than reasoning on the basis of non-crystallized deductive strategies (see also Anderson, 1982).

The working memory model proposed by Baddeley and Hitch (1974) and elaborated by Baddeley (1986, 1992, 1997) provides a suitable framework for the investigation of working memory involvement in cognitive performance. The model distinguishes a central executive from two subordinate memory systems, that is, the phonological loop for the retention of verbally coded material and the visuospatial sketchpad for the retention of visually or spatially coded information. The working memory model is closely tied to the dual task methodology, which allows one to examine which components of working memory contribute to the performance in a given task. In dual task experiments, a secondary task that loads the resources of a specific component of working memory has to be carried out concurrently to the criterion task. The degree of disruption under dual task conditions, as compared with a single task condition, is taken to reflect the dependence of the criterion task on the particular working memory component.

Hitherto, studies applying the dual task methodology in the domain of propositional reasoning have brought inconsistent results (see Gilhooly, 1998). In an experiment by Evans and Brooks (1981), neither a concurrent articulation task nor a concurrent retention task interfered with accuracy in conditional reasoning. Thus, this study provided no hint at an involvement of the phonological loop or the central executive in propositional reasoning. Toms, Morris, and Ward (1993), however, found that the central executive, but neither the phonological loop nor the visuospatial sketchpad was involved in conditional reasoning. The same pattern of results was reported by Gilhooly, Logie, Wetherick, and Wynn (1993) in the domain of syllogistic reasoning. Using a propositional inference task with various logical connectives, Klauer, Stegmaier, and Meiser (1997) also found that concurrent loads
of the phonological loop and the visuospatial sketchpad had no significant effects on the accuracy of inferences, although slight deteriorations were observed in terms of response latencies and secondary task performance. In addition, Klauer et al. replicated the disruptive effect of a central executive load on reasoning performance.

Taken together, the majority of results suggest that propositional reasoning relies on the resources of the central executive to some extent, but that it appears uncertain at best whether reasoning performance also depends on the subordinate working memory systems. The latter is surprising insofar as propositional reasoning entails verbal processing and memory, such as language comprehension and storage of verbal premises and assertions, so that it can be surmised that the phonological loop is also involved. Moreover, according to the mental model theory, mental models are quasi-analog representations of possible states of events (Johnson-Laird, 1983, 1996), and the construction and storage of those representations might additionally draw on the resources of the visuospatial sketchpad.

1.3. The moderating role of processing systems

The weak effects of secondary tasks on propositional reasoning may in part be due to the function of non-analytic, heuristic processes in reasoning. In fact, current dual process models of thinking emphasize the moderating role of analytic versus non-analytic strategies for reasoning performance as well as for the mental resources involved. For example, the two-stage theory of human inference (Evans, 1984, 1989) states that pre-attentive heuristics of information selection precede analytic processes of deductive inference. In the case of biased heuristic information selection, even perfect analytic inference processes may lead to incorrect conclusions. The more recent dual process theory proposed by Evans and Over (1996) distinguishes reasoning according to a personal system of goals and beliefs, called rationality1, from reasoning according to a normative system of logic, called rationality2. Whereas rationality1 is dominated by implicit and tacit mental processes, rationality2 is based on explicit and conscious processes of deductive inference. Similar distinctions were proposed by Sloman (1996) and by Stanovich (1999). According to Sloman (1996), an associative system of human thought is based on perceived correlations and temporal contiguities between concrete concepts, and a rule-based system is based on symbolic manipulations of abstract tokens. Stanovich (1999) delineates a dual process model comprising a system of automatic and largely unconscious mental processes, called System 1, and a system of controlled and analytic reasoning strategies, called System 2. The processes of System 1 serve evolutionary goals and impose low demands on individuals’ cognitive resources. System 1 reasoning is further characterized by contextualized task construals that often differ from the interpretation of formal logic. In contrast, the processes of System 2 serve individual goals and require higher amounts of cognitive capacity and mental resources. In this system, tasks are construed in an abstract manner, allowing for deductive inference in line with formal logic.

Following the dual process accounts, concurrent loads of working memory are expected to disrupt the analytic reasoning processes of rationality2 or System 2, while
leaving the heuristic processes of rationality\textsubscript{1} or System 1 largely unaffected. Hence, if participants employ heuristic processes to a large extent, dual task experiments on propositional reasoning may essentially underestimate the role of working memory in deductive inference. Thus, the inconsistent and mild interference effects observed so far may be accounted for in terms of the function of non-analytic processes that limit susceptibility of reasoning performance to secondary task disruption.

1.4. Training propositional reasoning

If working memory involvement in propositional reasoning is diluted by the use of non-analytic processes, how can the deployment of cognitive resources in “truly” deductive thinking be analyzed? Gilhooly et al. (1993) suggested that sufficient training prior to dual task experiments may reduce the impact of heuristics and strengthen the use of analytic reasoning strategies (see also Gilhooly, Logie, & Wynn, 1999). In fact, there is evidence that training cognitive skills affects working memory involvement and can lead to more differentiated patterns of secondary task interference (Logie, 1995; Logie, Baddeley, Mané, Donchin, & Sheptak, 1989). Applied to the case of propositional reasoning, training can inhibit the function of heuristics (see Houdé & Moutier, 1996) and encourage participants to use logically valid task construals and analytic inference processes, for example by pointing out contradictions between subjective inferences and normatively correct responses (see Klaczynski & Laipple, 1993; O’Brien & Overton, 1980). Pretraining may therefore be an experimental method to investigate working memory involvement in analytic reasoning processes pertaining to rationality\textsubscript{2} or System 2.

Following this rationale, Klauer et al. (1997, Experiment 4) ran training sessions prior to a dual task experiment on visuospatial interference in propositional reasoning. During the pretraining phase, comprehension of logical connectives was exercised by means of a truth table task with immediate feedback. In the subsequent dual task experiment, a secondary task loading the visuospatial sketchpad caused substantial disruption in reasoning accuracy. Moreover, disruption was significantly more pronounced after the comprehension training than in a control group without prior training, in which no disruption occurred. The results suggested that analytic reasoning processes, provoked by the training, are susceptible to visuospatial interference, and that this effect may be concealed by non-analytic processes in conditions without prior training.

Since the mental model theory of reasoning bears a strong resemblance to the definition of premises in terms of truth tables, the pretraining in the Klauer et al. (1997) study practiced analytic reasoning mainly by the use of mental models. That is, exercising truth tables for the premises of the reasoning task may facilitate the construction of appropriate initial models, improve the derivation of putative conclusions, and enhance the ability to flesh out models in the search for counter-examples. In contrast, mental logic theories and domain-specific reasoning theories imply different sorts of training to be most effective in reducing heuristics and in fostering analytic inference processes. Mental logic theories predict that a training condition which clarifies the syntactic structure of the premises and exercises the
application of inference rules should increase the accessibility of such rules and facilitate their use in deduction (Smith, Langston, & Nisbett, 1993). According to the theory of pragmatic reasoning schemata, one would predict optimal benefit from a treatment which provides a thematic context evoking appropriate pragmatic schemata for the premises at hand (Cheng & Holyoak, 1985; Cheng et al., 1986).

In a recent training evaluation study, Klauer, Meiser, and Naumer (2000) compared the effects of different training conditions that were based on the mental model theory, mental logic theories, and domain-specific accounts of reasoning. Truth table trainings served to enhance reasoning processes by mental models, rule-based trainings practiced the derivation of proofs and conclusions on the basis of mental logic theories, and domain-specific training conditions provided a pragmatic context to facilitate inferences from general knowledge structures. In two experiments with varying operationalizations of the different sorts of training, the truth table trainings and the domain-specific trainings were equally effective in improving reasoning performance. The training effects of these two conditions exceeded those of the rule-based trainings and of a pseudo-training serving as a baseline condition.

1.5. Goals and hypotheses

In the present experiments, the training conditions designed and evaluated by Klauer et al. (2000, Experiment 2) were used to extend the investigation of working memory involvement in analytic reasoning processes. For this purpose, the experiments addressed interference in the visuospatial sketchpad (Experiment 1), the central executive (Experiment 2), and the phonological loop (Experiment 3) after different kinds of pretraining. The analysis of interactions between secondary task effects and training conditions allowed us to test whether analytic reasoning processes, pertaining to rationality or System 2, were differentially evoked by the training conditions. According to the mental logic view, a rule-based training should enhance analytic reasoning processes using natural inference rules, thereby increasing vulnerability to concurrent loads of working memory. Following the mental model theory, increased susceptibility to secondary task interference is predicted after the truth table training, which aims to strengthen analytic reasoning by mental models. Since the theory of pragmatic reasoning schemata implies that schema activation releases from demanding reasoning processes, interference is expected to decrease, rather than increase, as a consequence of the domain-specific training.

2. Experiment 1

The first experiment addressed interference in propositional reasoning performance due to a concurrent spatial tapping task. Spatial tapping has proved to cause modality-specific visuospatial interference in working memory (Farmer, Berman, & Fletcher, 1986; Meiser & Klauer, 1999), and it has been used in numerous studies to investigate the involvement of the visuospatial sketchpad in cognitive tasks (e.g., Gilhooly et al., 1993; Klauer et al., 1997; Morris, 1987; Toms et al., 1993). In
particular, the first experiment aimed to extend previous findings by Klauer et al. (1997, Experiment 4) on visuospatial interference in propositional reasoning following a pretraining session. As mentioned above, Klauer et al. demonstrated substantial disruption by concurrent spatial tapping after a truth table training. The authors interpreted the visuospatial interference as corroborating the representational assumptions of the mental model theory in terms of quasi-analog or isomorphic representations of possible states of the world (see Johnson-Laird, 1983, 1996). In the training phase of this previous study, a truth table task was employed and, in the case of a wrong response, the meaning of the premises was clarified by free discussion with the experimenter. In the present experiment, different training conditions were realized, and all explanatory texts were provided in a fully standardized format, thereby minimizing the potential influence of demand characteristics. The training conditions were taken from the aforementioned training evaluation study (Klauer et al., 2000).

2.1. Method

2.1.1. Participants
Eighty undergraduate students from various departments of the University of Bonn volunteered in the experiment and received 10 DM for their participation. The participants were quasi-randomly assigned to four training conditions of equal sample size.

2.1.2. Materials
Reasoning task. The premises of the reasoning task were presented acoustically using a personal computer equipped with a sound card and headphones. In each trial of the reasoning task, participants heard two premises. The first one was a composite premise of the kind $pCq$, where $C$ denotes one of four logical connectives:

- “If $p$ then $q$.” (Material implication),
- “$p$ if and only if $q$.” (Equivalence),
- “$p$ and/or $q$.” (Inclusive disjunction),
- “Either $p$ or $q$.” (Exclusive disjunction).

The symbols $p$ and $q$ indicate elementary propositions that could be affirmative or negated. The elementary propositions referred to the presence of arbitrary geometrical objects in the case of affirmative propositions and to the absence of such objects in the case of negated propositions. Examples of the affirmative and negated elementary propositions are “there is a circle”, “there is a triangle”, “there is no cube”, and “there is no ellipse”. In the domain-specific training condition, the geometrical objects were replaced by objects and events that had a pragmatic meaning within the context established by the pretraining (see below).

The second premise of a trial was a minor premise concerning one of the propositions $p$ and $q$. The minor premise represented one of the four logical cases $p$, $\neg p$ (i.e., the complement of $p$), $q$, and $\neg q$ (i.e., the complement of $q$). Participants had to decide what follows from the two premises with respect to the object not mentioned in the minor premise. In each trial, three response alternatives were available, namely
(1) that the object is there, (2) that the object is not there, and (3) that nothing follows. After having formed a conclusion, participants pressed the spacebar on the computer keyboard, whereupon the three response alternatives were displayed on the screen and a response interval was initiated. Participants responded orally by reading off the selected alternative into a microphone. Table 1 shows the correct response alternative for each connective of the composite premise and logical case of the minor premise depending on the affirmation and negation of the propositions $p$ and $q$.

To assess secondary task interference in a within-participants design, each participant worked through two blocks with 16 trials of the reasoning task, that is, one block with and the other block without concurrent tapping. The 16 trials of a block were generated by crossing the four logical connectives by the four logical cases of the propositions $p$ and $q$.

### Table 1
Correct responses in the propositional reasoning task with affirmative and negated propositions

<table>
<thead>
<tr>
<th>Minor premise</th>
<th>Composite premise</th>
<th>$pCq$</th>
<th>$pCq$</th>
<th>$pCq$</th>
<th>$pCq$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$++$</td>
<td>$+-$</td>
<td>$-+$</td>
<td>$--$</td>
</tr>
<tr>
<td>'If $p$ then $q$'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>$1^b$</td>
<td>2$^c$</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$\neg p$</td>
<td>3$^d$</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>$q$</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>$\neg q$</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>'If $p$ and only if $q$'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$\neg p$</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$q$</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$\neg q$</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>'If $p$ or $q$'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>$\neg p$</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$q$</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>$\neg q$</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>'Either $p$ or $q$'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\neg p$</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$q$</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$\neg q$</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Notes: + indicates an affirmative proposition (e.g., “there is a circle”); – indicates a negated proposition (e.g., “there is no cube”). The response alternatives displayed in the body of the table refer to the presence or absence of the object not mentioned in the minor premise.

$^b$ The object is there.

$^c$ The object is not there.

$^d$ Nothing follows.
the minor premise so that each trial corresponded to one row in Table 1. In constructing the first block, affirmation and negation of the propositions $p$ and $q$ were determined for each trial by random selection from the four possible combinations (see columns of Table 1). The trials of the second block were composed by random selection from the remaining three combinations. The assignment of affirmation and negation to the trials of the two blocks was held constant within pairs of participants and counterbalanced across secondary task condition. The order of the 16 trials was randomized per block and participant.

*Training conditions.* The four training conditions included a truth table training, a rule-based training, a domain-specific training, and a pseudo-training that served as baseline. All training conditions were computerized, that is, the problems and explanatory texts were presented visually on the computer screen, and the responses were entered manually on the keyboard or with the computer mouse. The pre-training phase consisted of two training blocks of 16 trials. Participants received automatic feedback on the correctness of their responses within each trial. After each training block, a summary of the performance in the last 16 trials was provided. After the second training block, participants were also informed about their improvement relative to the first block.

In each trial of the *truth table training*, a composite premise of the kind $pCq$ was presented together with the four combinations of true and false propositions generating the truth table. The participants were to assign one of the truth values “true” and “false” to each of the combinations. A training block comprised four trials for each of the four connectives, with complete permutation of affirmation and negation of the propositions $p$ and $q$. Upon completion of the truth table for a given composite premise, participants received immediate feedback about the accuracy of their responses and about the correct truth values. During the first training block, participants were presented with an explanation of the meaning of the composite premise whenever they made an error. The explanation contained a verbal description of the connective and an abstract definition in terms of a truth table. Participants were invited to insert the propositions of the current premise into the abstract table. Finally, the trial was re-presented together with the participants’ responses and the correct truth values. During the second training block, participants were offered the explanatory text in the case of an error, but they could skip the text if they wished.

In the rule-based training condition, participants received a set of premises and a conclusion for each trial. They had to decide whether the conclusion is true or false on the basis of the premises. A pool of 32 trials was created from the problems used by Braine et al. (1984, 1995). Deviating from the original problems, the premises were phrased using affirmative and negated propositions about the same geometrical objects as in the reasoning task. Furthermore, some problems were altered by

---

1 A complete documentation of the explanatory texts for examples of the premises is available in an online appendix to Klauer et al. (2000, Experiment 2), via internet: http://www.psychologie.uni-bonn.de/sozial/forsch/lsf/E2as/e2asht.htm#p1.
introducing the conjunctions “if and only if” and “either or”, which did not occur in
the original problems, so that all relevant connectives were exercised. For each
participant, the 32 trials were quasi-randomly divided into two equivalent blocks.
During the first training block, an explanatory text was displayed after each re-
sponse. The text demonstrated the sequence of natural inference rules that lead to the
correct conclusion for the given problem as specified by Braine et al. (1984, 1995).
For premises with connectives “if and only if” and “either or”, the sequences of
inference rules were adapted accordingly. The explanations clarified the syntactic
structure of the premises and elucidated the application of each inference rule in the
sequence. While the inference steps were described one by one, the actual problem as
well as intermediate results of the deduction chain remained visible on the screen. 2
Following the explanatory text, participants were informed whether their response to
the problem had been correct or not. The requirement to read the explanations to all
problems in the first block, irrespective of the participants’ responses, was to guar-
anteed a sufficient “dose” of rule-based training instruction despite the 50% guessing
rate of correct responses. During the second training block, participants received
immediate feedback about the correctness of their decisions, and they could choose
whether or not the explanatory texts should be presented.

In the domain-specific training condition, the training phase started with an in-
trductory text that provided comprehensive background information about a fic-
titious American company. The text described a set of rules which are in force to
regulate working hours and off-times in the company (e.g., an employee takes a day
off during the week if and only if he has previously worked on a weekend), to prevent
accidents and job-related diseases (e.g., the company is not allowed to have an
employee work in a particular factory zone if he shows a disposition to develop
certain diseases), and to maintain the quality of the manufactured products (e.g., a
certain bonus is awarded if and only if the quality of products in a given segment has
not dropped). Each rule was delineated in a separate paragraph which emphasized
the relevance of the rule and explained its necessity and sufficiency relations in terms
of the company regulations. The set of rules was designed according to the theory of
pragmatic reasoning schemata and included a precautionary permission rule (Cheng
& Holyoak, 1989), a social obligation rule (Cheng et al., 1986), a precautionary
obligation rule (Cheng & Holyoak, 1989), and two pragmatically biconditional rules
comprising a permission and a complementary obligation (Holyoak & Cheng, 1995).
The rules described in the introductory text matched the composite premises of the
reasoning task and thus established a context in which the domain-specific inference
strategies of the permission and the obligation schemata should become available.
After reading the introductory text, participants worked through two blocks of 16
questions about the company. Each question consisted of a statement related to one
of the rules, and the participants had to decide whether the statement applies or not.
They received immediate feedback about the correctness of their response. When an

2 A complete documentation of the 32 problems and the corresponding explanatory texts are available
erroneous response occurred during the first training block, a paragraph of the introductory text was re-presented in which the rule referring to the statement was described. Finally, the statement was displayed again together with the correct response and a sentence summarizing the essence of the paragraph for the statement at hand. When an erroneous response occurred during the second training block, participants were offered the explanation with the option to skip it.

For the pseudo-training condition, 16 types of inductive analogy problems were created which resembled typical problems in Raven’s (1962) progressive matrices test. Each problem consisted of a $3 \times 3$ matrix with eight figures and one empty cell. In half of the problems the figures were composed of lines and shapes, in the other half of nested textures. The matrices were constructed according to the principles underlying most of the Raven problems, such as distribution of three values, invariance of attributes within rows or columns, or quantitative pairwise progression, as identified by Carpenter, Just, and Shell (1990). The participants worked through two training blocks with varying realizations of the 16 types of inductive problems. They had to select the missing matrix figure out of six alternatives. When a wrong alternative was chosen during the first training block, an explanatory text was displayed which elaborated on the principles underlying the present matrix. The explanation described the structure of the figures and emphasized the need to infer the construction principles of a given matrix to reach the correct response. Following the text, the matrix of the problem was re-presented with the missing figure highlighted by a coloured frame. When a wrong alternative was chosen during the second training block, the explanatory text was offered with the option to skip it.

Spatial tapping. In the secondary task of Experiment 1, the participants had to tap the fields of a $5 \times 5$ array in a fixed spatial sequence at a rate of one field per second. The $5 \times 5$ array was fitted on a square board with a side length of 36 cm. The board was concealed from vision to avoid visual distraction and to increase spatial monitoring while performing the task. The tapping task was practiced for one minute, before it was combined with the reasoning task.

2.1.3. Procedure

The participants went through the procedure individually. At the beginning of an experimental session, participants received a written instruction about the reasoning task, including a brief introduction into the meaning of the conjunctions “if then”, “if and only if”, “and/or”, and “either or”. It was emphasized that “$p$ if and only if $q$” denotes a biconditional statement, corresponding to “If $p$ then $q$, and if $q$ then $p$”. It was also stressed that “$p$ and/or $q$” denotes an inclusive disjunction, that is, “$p$ or $q$, or both”, and that “either $p$ or $q$” denotes an exclusive disjunction, that is, “$p$ or $q$, and not both”. These interpretations conform to the use of the conjunctions in colloquial German. After the general instruction, participants were familiarized with

---

3 A complete documentation of the paragraphs describing the rules and of the 32 statements and answers are available via internet: http://www.psychologie.uni-bonn.de/sozial/forsch/lsf/E2In/E2inho.htm#p1.
the reasoning task by eight example trials. Then they worked through one of the four kinds of pretraining. Following the training phase, the two experimental blocks with 16 trials of the reasoning task were administered. In one of these blocks, participants concentrated on the reasoning task alone. In the other block, they had to perform the tapping task concurrently to the reasoning task throughout the 16 trials of this block. The order of the two blocks as well as the random assignment of affirmative and negated propositions to the two blocks were counterbalanced across participants within each training condition.

2.2. Results

The mean number of correct responses in the reasoning task is displayed in Table 2 as a function of training condition and concurrent tapping. Since the secondary task was expected to reduce, rather than improve, reasoning performance, the statistical test of the main effect of concurrent tapping was one-tailed. No a priori hypotheses were specified for the main effect of training condition, because the experimental design did not allow for comparisons of training gains (see Klauer et al., 2000, for a pre-post evaluation of training conditions). For all tests reported in this article, the level of statistical significance was set to $\alpha = 0.05$.

A 4 (training condition) $\times$ 2 (tapping) analysis of variance with repeated measures on the last factor revealed a significant main effect of training condition, $F(3, 76) = 7.89, P < 0.001$, and of concurrent tapping, $F(1, 76) = 5.03, P = 0.014$ one-tailed. The effect of concurrent tapping was not moderated by an interaction with training condition, $F(3, 76) < 1$. Closer inspection of the main effect of training condition using Tukey’s honestly significant difference (HSD) tests showed that reasoning performance was lower in the pseudo-training condition than in each of the remaining three conditions. No further difference reached significance.

An additional analysis of variance including the factor of logical connective yielded a significant main effect of connective, $F(3, 228) = 55.11, P < 0.001$, but no interaction of this factor with training condition, $F(9, 228) = 1.56, P = 0.128$, or concurrent tapping, $F(3, 228) < 1$, and no three-way interaction, $F(9, 228) < 1$. Overall, premises expressing an equivalence elicited the largest number of correct

<table>
<thead>
<tr>
<th>Training condition</th>
<th>Without tapping</th>
<th>Concurrent tapping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>SD</td>
</tr>
<tr>
<td>Truth table training</td>
<td>9.90</td>
<td>2.71</td>
</tr>
<tr>
<td>Rule-based training</td>
<td>10.05</td>
<td>2.56</td>
</tr>
<tr>
<td>Domain-specific training</td>
<td>11.75</td>
<td>2.29</td>
</tr>
<tr>
<td>Pseudo-training</td>
<td>8.60</td>
<td>2.39</td>
</tr>
</tbody>
</table>

*Note: Maximum number of correct responses = 16.*
responses ($M = 6.33$, $SD = 1.46$; maximum number = 8), premises expressing an inclusive disjunction the lowest number ($M = 3.98$, $SD = 1.79$), with exclusive disjunctions ($M = 4.85$, $SD = 1.98$) and material implications ($M = 4.38$, $SD = 1.38$) in between.

2.3. Discussion

Experiment 1 showed significant disruption in propositional reasoning performance due to a spatial tapping task. In contrast to the previous findings by Klauer et al. (1997, Experiment 4), the disruptive effect was not moderated by the kind of pretraining in the present data. This discrepancy may be explained by procedural differences, such as the use of a no-training baseline condition in the earlier study versus a pseudo-training condition that closely resembled the actual trainings in structural and motivational aspects in the present experiment. However, the present experiment gave no indication whatsoever of an increased, or decreased, susceptibility of reasoning performance to visuospatial interference following any of the training conditions. Hence, the results provided no evidence that the presumed enhancement of analytic reasoning processes coincides with a stronger involvement of the visuospatial sketchpad, nor that the activation of pragmatic schemata leads to a decreased involvement of this working memory component.

3. Experiment 2

The second experiment addressed the role of the central executive in propositional reasoning after the different kinds of training. To impose a load on the central executive, we used a secondary task that required the generation of random sequences of numbers. According to the working memory model, the generation of random sequences draws on the control functions of the central executive to supervise the process of random selection and to suppress stereotypic response patterns (Baddeley, 1986, 1992, 1996; Baddeley, Emslie, Kolodny, & Duncan, 1998). Following this rationale, random number generation has been used as a secondary task to load the central executive in various areas of research, such as syllogistic reasoning (Gilhooly et al., 1993), propositional reasoning (Klauer et al., 1997), mental arithmetic (Logie, Gilhooly, & Wynn, 1994), mental rotation (Logie & Salway, 1990), and serial short-term memory (Meiser & Klauer, 1999; Salway & Logie, 1995). In the present experiment, disruption of propositional reasoning by concurrent random number generation was analyzed using a control condition of concurrent counting. Since the verbal and articulatory demands of the two secondary tasks were approximately equal, disruptive effects of random generation as compared with concurrent counting can be attributed to the specific load that the random generation task imposes on the central executive. Differences in the amount of disruption between training conditions thus reflect differential involvement of the central executive.
3.1. Method

3.1.1. Participants

Eighty participants, most of them undergraduate students from various departments of the University of Bonn, volunteered in the experiment. The participants were quasi-randomly assigned to the four training conditions of equal sample size.

3.1.2. Materials

Reasoning task. The reasoning task was the same as in Experiment 1 except for the mode of presentation and responding. Because of the articulatory secondary tasks, the premises were now presented visually, and the responses were entered manually, to avoid structural interference in perception and reaction (Kahneman, 1973). In each trial of the reasoning task, the composite premise was displayed on the computer screen for 8.2 s. Following a pause of 0.3 s, the minor premise was displayed for 4.2 s. After having formed a conclusion, the participants initiated the menu of response alternatives (see Section 2.1) and entered the value of their response on the keyboard. The successive presentation of premises and the choice of responses thus followed the procedure of the acoustical reasoning task in the previous experiment.

Training conditions. The four training conditions described in Section 2.1 were used during the training phase of the present experiment.

Secondary tasks. In the random number generation task, participants had to pronounce digits from 1 to 9 in random order. Random generation was explained as random selection with replacement and illustrated by the urne metaphor. In the counting condition, participants were asked to count repeatedly from 1 to 9. In both tasks, the digits had to be pronounced at a rate of one digit per second throughout the 16 trials of a reasoning block. The secondary tasks were practiced for one minute, before they were combined with the reasoning task. The articulated digits were recorded by means of a tape recorder. Following the experimental session, the random digits were entered into the computer for statistical analysis.

3.1.3. Procedure

The procedure of Experiment 2 followed that of the previous experiment. After the training phase, the participants worked through two blocks of the reasoning task, one block with concurrent random number generation and one block with concurrent counting. The order of the two blocks as well as the random assignment of affirmative and negated propositions to the two blocks was counterbalanced within each training condition.

3.2. Results

Table 3 displays the mean number of correct responses in the reasoning task for the different training conditions and secondary tasks. Since random number generation was expected to yield a larger, rather than smaller, disruptive effect on reasoning performance than concurrent counting, the test for the secondary task effect was one-tailed. No a priori hypotheses were specified for the main effect of training condition.
In a 4 training condition × 2 (secondary task) analysis of variance with repeated measures on the last factor, the main effect of training condition did not attain significance, $F(3, 76) = 1.51, P = 0.218$. However, the effect of secondary task was significant, $F(1, 76) = 8.56, P = 0.003$ one-tailed, and the interaction of training condition and secondary task approached significance, $F(3, 76) = 2.65, P = 0.055$. As Table 3 indicates, the interaction results from an especially marked disruption of reasoning performance by random number generation in the truth table training condition. In fact, a decomposition of the interaction using contrast analysis (Rosenthal & Rosnow, 1985) revealed that random number generation caused larger interference in the truth table training condition than in the other three conditions, $F(1, 76) = 6.86, P = 0.011$, while the degree of interference did not differ among the latter three, $F(2, 76) < 1$.

An additional analysis of variance with the factor of logical connective showed a significant effect of connective, $F(3, 228) = 18.28, P < 0.001$, but no interaction with training condition, $F(9, 228) < 1$, or secondary task, $F(3, 228) = 2.07, P = 0.105$, and no three-way interaction, $F(9, 228) = 1.04, P = 0.405$. Overall, premises expressing an equivalence yielded the largest number of correct conclusions ($M = 4.70, SD = 1.68$; maximum number = 8), premises expressing an inclusive disjunction the lowest ($M = 3.00, SD = 1.63$), with exclusive disjunctions ($M = 3.79, SD = 1.67$) and material implications ($M = 3.43, SD = 1.54$) in between.

To test for differences in random generation performance between training conditions, one-way analyses of variance were conducted on the number of generated random digits and on the first-order redundancy of the generated sequences. First-order redundancy reflects deviations of the generated sequences from a uniform distribution over the nine digits and thus is a measure of randomness (Attnaveve, 1959). The analyses yielded neither differences in the number of generated digits, $F(3, 76) = 1.35, P = 0.265$, nor in redundancy, $F(3, 76) < 1$.

### 3.3. Discussion

Experiment 2 showed larger disruption of reasoning performance by concurrent random generation after the truth table training than after the other kinds of
training. Since the experimental design controlled for verbal and articulatory demands, this result can be interpreted to reveal enhanced involvement of the central executive in propositional reasoning in the truth table training condition. Because no differences were obtained in the measures of random generation performance as a function of training condition, the larger disruptive effect following the truth table training cannot be traced back to a differential trade-off between reasoning and secondary task. To rule out another alternative explanation, namely that the larger disruption in the truth table training condition might be an artifact of higher baseline performance (cf. Loftus, 1978), the above contrast analysis was repeated with a transformed measure of interference. In the new analysis, the dependent variable was the percentage decrement in reasoning performance due to random number generation relative to the baseline performance under concurrent counting. This transformation standardizes the amount of disruption by the level of baseline performance and therefore diminishes the risk of artifactually increased interference (see Logie & Salway, 1990; Meiser & Klauer, 1999). The new contrast analysis brought the same results as the previous analysis of the untransformed difference score: Random number generation was more disruptive following the truth table training than following the other kinds of training, $F(1, 76) = 4.89, p = 0.030$, whereas no differences occurred between the latter three, $F(2, 76) < 1$. This outcome confirmed the conjecture that propositional reasoning depends on the central executive to a larger extent after the truth table training than after the other kinds of training, and it safeguarded the interpretation against possible scale artifacts. We can conclude that, among the training conditions used in the present study, the truth table training emerged to be unique in eliciting reasoning processes that are particularly vulnerable to limitations of central executive resources.

4. Experiment 3

Experiment 3 investigated the role of the phonological loop in propositional reasoning after the different kinds of training. Concurrent articulation causes modality-specific interference (Farmer et al., 1986; Meiser & Klauer, 1999) and has therefore been used in numerous studies to investigate the dependence of cognitive tasks on the phonological loop (e.g., Gilhooly et al., 1993; Morris, 1987; Toms et al., 1993). In the previous experiment, concurrent counting was used as a control condition to cancel out the effects of verbal and articulatory demands from the effects of random number generation. The present experiment targeted directly on the effects of an articulatory load on reasoning performance by comparing a counting condition with a control condition of no concurrent load. As mentioned above, earlier studies had shown negligible effects of concurrent articulation on reasoning performance (Evans & Brooks, 1981; Gilhooly et al., 1993; Klauer et al., 1997; Toms et al., 1993), indicating only minor involvement of the phonological loop. The goal of Experiment 3 therefore was to examine whether any of the training conditions renders a noticeable susceptibility of reasoning performance to articulatory interference.
4.1. Method

4.1.1. Participants

Eighty undergraduate students from various departments of the University of Bonn volunteered in the experiment and received 10 DM for their participation. The participants were quasi-randomly assigned to the four training conditions of equal sample size.

4.1.2. Materials

Reasoning task. The reasoning task was administered in the visual presentation mode with manual responses, as described in Section 3.1.

Training conditions. The four training conditions described in Section 2.1 were used during the training phase of the present experiment.

Counting. Articulatory suppression by counting was used as secondary task in Experiment 3. As in the previous experiment, participants had to count repeatedly from 1 to 9 at a rate of one digit per second throughout the 16 trials of a reasoning block. The secondary task was practiced for one minute, before it was combined with the reasoning task.

4.1.3. Procedure

The procedure of Experiment 3 followed that of Experiments 1 and 2. After the training phase, participants worked through two blocks of the reasoning task, one block without secondary task and one block with concurrent counting. The order of the two blocks as well as the random assignment of affirmative and negated propositions to the two blocks were counterbalanced within each training condition.

4.2. Results

Table 4 displays the mean number of correct responses in the reasoning task as a function of training condition and concurrent counting. Since the secondary task was expected to reduce, rather than enhance, reasoning performance, the corresponding statistical test was one-tailed. No a priori hypotheses were specified for the main effect of training condition.

A 4 (training condition) × 2 (counting) analysis of variance with repeated measures on the last factor yielded no significant main effect of training condition, \( F(3,76) < 1 \), but a significant effect of concurrent counting, \( F(1,76) = 7.28, P = 0.005 \) one-tailed. Although the overall interaction was not significant, \( F(3,76) = 1.18, P = 0.324 \), a contrast analysis using the same contrasts as in the previous experiment revealed a strong tendency for the disruptive effect to be more pronounced in the truth table condition than in the other three conditions, \( F(1,76) = 3.32, P = 0.072 \), without differences among the latter, \( F(2,76) < 1 \).

An additional analysis of variance with the factor of logical connective showed a significant effect of connective, \( F(3,228) = 40.18, P < 0.001 \), and an interaction of connective with concurrent counting, \( F(3,228) = 4.82, P = 0.003 \). The interaction with training condition and the three-way interaction were not significant,
F(9, 228) = 1.71, P = 0.089, and F(9, 228) < 1. In both blocks of the reasoning task, premises expressing an equivalence were least difficult (without counting: M = 3.23, SD = 0.81; with counting: M = 2.79, SD = 0.99; maximum number of correct responses = 4), and premises expressing an inclusive disjunction most difficult (M = 1.91, SD = 1.20; M = 1.88, SD = 1.04), with material implications (M = 2.29, SD = 0.78; M = 1.90, SD = 1.06) and exclusive disjunctions (M = 2.24, SD = 1.17; M = 2.36, SD = 1.01) in between.

4.3. Discussion

Unlike several previous studies (Evans & Brooks, 1981; Gilhooly et al., 1993; Klauer et al., 1997; Toms et al., 1993), the present experiment demonstrated appreciable disruption of reasoning accuracy by articulatory suppression. Although only marginally significant, the contrast analysis indicated that the disruption was more pronounced in the truth table training condition than in the other training conditions. In fact, the disruptive effect of articulatory suppression was largely confined to the truth table condition, as shown in Table 4 and supported by post hoc t-tests: The interference effect was significant in the truth table training condition, t(19) = 2.49, P = 0.011 one-tailed, but it did not attain significance in the rule-based, t(19) = 1.21, P = 0.122 one-tailed, domain-specific, t(19) < 1, and pseudo-training, t(19) < 1, condition. That is, with the exception of the truth table training condition, Experiment 3 replicated the null effects found in earlier studies. The results therefore indicate that the truth table training provokes reasoning processes which are dependent on the availability of resources of the phonological loop.

5. General discussion

The starting point of the present experiments was the observation that previous studies had yielded inconsistent findings concerning the role of working memory in propositional reasoning, with generally weak interference by secondary tasks loading the subcomponents of working memory. Taking into account the moderating effects of analytic and non-analytic processing systems on task performance and cognitive
resources, which have been highlighted by current dual process accounts of reasoning (e.g., Evans & Over, 1996; Sloman, 1996; Stanovich, 1999), we assumed that reasoning performance reveals little dependence on working memory due to the overpowering function of non-analytic heuristic processes. As non-analytic processes are less cognitively demanding than analytic processes of reasoning (see Stanovich, 1999), previous dual task experiments may have underestimated working memory involvement in human deduction.

In the present experiments, the role of non-analytic processes in task construal and deductive inference was tackled by means of different pretraining conditions. The training conditions were designed on the basis of a mental logic theory, the mental model theory, and the theory of pragmatic reasoning schemata. That is, they aimed to enhance the application of mental inference rules, to strengthen model-based reasoning processes, or to trigger facilitative reasoning schemata. The analysis of secondary task effects as a function of prior training allowed us to test competing predictions concerning working memory involvement in analytic reasoning processes. Mental logic theories and the mental model theory imply that analytic reasoning by inference rules and mental models, respectively, is particularly dependent on working memory resources, so that the training conditions related to these theories should increase susceptibility to secondary task interference. The theory of pragmatic reasoning schemata, in contrast, contends that schema activation reduces the cognitive demands of inference processes and thus decreases interference.

The major results were as follows. Experiment 1 replicated the disruption of propositional reasoning by concurrent spatial tapping reported by Klauer et al. (1997, Experiment 4). In contrast to the earlier study, however, the extent of spatial interference was not moderated by the kind of pretraining. Experiments 2 and 3 revealed that random number generation and concurrent articulation were particularly disruptive after the truth table training as compared with the other kinds of training. The experiments thus indicated that, among the training conditions realized here, the truth table training rendered particular susceptibility to concurrent loads of the central executive and the phonological loop.

Before considering the implications of our findings, two alternative interpretations of the data have to be addressed. Both alternative interpretations question our contention that the results reflect differential involvement of working memory components in the truth table training condition.

The first alternative interpretation is based on the observation that baseline reasoning performance appeared to be higher following the truth table training than following the other kinds of training in the Experiments 2 and 3 (see Tables 3 and 4). That is, the truth table training led to higher baseline performance than the

---

4 The main effect of training condition on baseline reasoning performance (i.e., reasoning performance under concurrent counting in Experiment 2 and without secondary task in Experiment 3) approached significance in Experiment 2, $F(3, 76) = 2.20, P = 0.095$, but not in Experiment 3, $F(3, 76) < 1$. Moreover, contrast analyses showed that performance was significantly higher after the truth table training than after the other kinds of training in Experiment 2, $F(1, 76) = 5.95, P = 0.017$, but not in Experiment 3, $F(1, 76) = 1.75, P = 0.190$. 

---
other training conditions in exactly those experiments that revealed larger interference for the truth table training group (see also Section 3.3). This observation suggests that the amount of interference is not a function of the type of secondary task used in the different experiments, but of the extent to which the truth table training induced cognitively demanding analytic reasoning processes that caused high task performance. Therefore, the increased secondary task effects after the truth table training in the Experiments 2 and 3 might not reflect specific involvement of the central executive and the phonological loop, but general susceptibility of analytic reasoning processes to any type of concurrent load.

To address this interpretation, we conducted analyses of variance of reasoning performance under baseline and secondary task conditions across the three experiments, that is, across the types of secondary tasks loading different components of working memory. If baseline performance in a given training condition is invariant across experiments and if secondary task interference varies as a function of experiment, then we may conclude that the amount of interference reflects differential involvement of working memory components. In fact, for the truth table training condition there was no effect of experiment on baseline performance, $F(2, 57) = 1.17, P = 0.317$. Most important, baseline performance in Experiment 1 did not differ from baseline performance in Experiments 2 and 3, $F(1, 57) < 1$. Thus, there was no indication that the truth table training was more successful in enhancing analytic reasoning processes in the latter experiments. However, in this training condition there was a marginally significant interaction between experiment and secondary task load, $F(2, 57) = 3.04, p = 0.055$, with significantly larger interference effects in Experiments 2 and 3 than in Experiment 1, $F(1, 57) = 5.64, p = 0.021$. Given invariance of baseline performance across experiments, this interaction corroborates the contention that reasoning performance in the truth table training condition specifically depended on the resources of the central executive and the phonological loop. No such interaction was found in the other training conditions, all $F(2, 57) < 1$. However, since baseline performance in these conditions did not prove invariant, the absence of an interaction is not stringently interpretable with respect to mental resources (see Loftus, 1978). Taken together, this post hoc analysis underpinned that amount of interference was a function of type of secondary task, rather than of training gain, in the truth table condition, while the results in the remaining training conditions were less clear-cut.

A second alternative interpretation of the data explains the disruptive effects by spatial tapping and articulatory suppression in terms of central executive demands, rather than in terms of interference in the modality-specific subcomponents of working memory. Inasmuch as the spatial tapping task and concurrent counting require streams of actions with changes between successive events, both secondary tasks may indeed cause amodal portions of interference (Jones, Farrand, Stuart, & Morris, 1995) that can be attributed to the function of the central executive (Meiser & Klauer, 1999). Although it cannot be ruled out that spatial tapping and concurrent

---

5 We are grateful to Jonathan Evans for pointing out this theoretical alternative.
counting load the central executive to some extent, the differential effects of the training conditions on secondary task interference render an interpretation in terms of a single working memory component unlikely. Given that the disruption by a secondary task loading the central executive was moderated by training condition in Experiment 2, the absence of any moderating effect of training condition on disruption by spatial tapping in Experiment 1 challenges the interpretation in terms of amodal interference located in the central executive. Moreover, the secondary tasks of spatial tapping and articulatory suppression have been shown to cause mainly modality-specific interference (e.g., Farmer et al., 1986; Meiser & Klauer, 1999; Morris, 1987), supporting the assumption that the disruptive effects in Experiments 1 and 3 primarily reflect interference in the visuospatial sketchpad and in the phonological loop.

How can the results reported herein be reconciled with the different theories of propositional reasoning outlined in the introduction? According to mental logic theories, the rule-based pretraining should enhance the accessibility and successful application of analytic inference processes that are vulnerable to limitations of working memory resources. Because the rule-based training condition did not affect susceptibility of reasoning performance to secondary task interference, however, the results do not provide support for the mental logic view. It remains possible, of course, that the expected effects might occur with a different rule-based training or with different secondary tasks. However, even if it was assumed that the rule-based training used here did not match the essence of reasoning processes based on natural inference rules or that the secondary tasks were not suitable to assess interference, mental logic theories could hardly account for the selective secondary task interference obtained in the truth table training condition.

The findings concerning the truth table training lend support to the mental model theory, which implies that exercising the comprehension of premises by means of truth tables strengthens analytic processes in the construction and selection of mental models (Klauer et al., 2000). The successful use of these processes is expected to be limited by working memory constraints, as borne out in the present experiments by the increased vulnerability to interference by random number generation and concurrent articulation following the truth table training. Moreover, the current results of the truth table training condition resemble recent findings in the domain of syllogistic reasoning. As demonstrated by Gilhooly et al. (1999), reasoning with abstract syllogisms following a comprehension training also shows remarkable susceptibility to concurrent loads of the central executive. The similarity of findings for propositional and syllogistic reasoning fits the notion of the mental model theory that analytic processes of model construction and selection are independent of the particularities of task structure and content.

Finally, the theory of pragmatic reasoning schemata gains partial support from the absence of any indication of increased secondary task interference following the domain-specific training. The facilitative effect of this training condition in the training evaluation study by Klauer et al. (2000) can be explained in terms of pragmatic schema evocation, causing a release from cognitively demanding reasoning processes. Providing the small overall amount of disruption in the above dual
task experiments, a demonstration of the expected decrease of interference may have been precluded by a floor effect. Under this post-hoc assumption, the absence of an increase in secondary task interference is in line with the theory of pragmatic schemata. Nonetheless, this theory disagrees with the findings obtained in the truth table training condition. As it is assumed that reasoners are not able to apply abstract knowledge about logical connectives to an inference task, unless they undergo a particular treatment focusing on the application procedure itself (Cheng et al., 1986), the abstract truth table training should not have affected performance and secondary task interference in the reasoning task, which it actually did.

In conclusion, the analysis of secondary task interference after different sorts of training yielded some support for the mental model theory, but challenged mental rule accounts and domain-specific explanations of propositional reasoning. Moreover, by highlighting the distinction between non-analytic heuristic strategies and cognitively demanding analytic processes, dual process accounts of human deduction have proved to be a suitable framework to obtain a more differentiated picture of working memory involvement in propositional reasoning.

References


