Figural Effects in a Syllogistic Evaluation Paradigm: An Inspection-Time Analysis

Edward J. N. Stupple

University of Derby, UK

Linden J. Ball*

Lancaster University, UK

*Corresponding author:

Department of Psychology, Lancaster University, Lancaster, LA1 4YF.

Tel: 01524 593470

Fax: 01524 593744

Email: L.Ball@lancaster.ac.uk

Abstract

Robust biases have been found in syllogistic reasoning that relate to the figure of premises and to the directionality of terms in given conclusions. Mental models theorists (e.g., Johnson-Laird & Byrne, 1991) have explained figural bias by assuming that reasoners can more readily form integrated models of premises when their middle terms are contiguous than when they are not. Biases associated with the direction of conclusion terms have been interpreted as reflecting a natural mode of reading off conclusions from models according to a "first-in, first-out principle". We report an experiment investigating the impact of systematic figural and conclusion-direction manipulations on the processing effort directed at syllogistic components, as indexed through a novel inspection-time method. The study yielded reliable support for mental-models predictions concerning the nature and locus of figural and directionality effects in syllogistic reasoning. We argue that other accounts of syllogistic reasoning seem less able to accommodate the full breadth of inspection-time findings observed.

Keywords

Syllogistic reasoning, figural effects, processing direction, inspection-time analysis, mental models, reasoning strategies

Figural Effects in a Syllogistic Evaluation Paradigm: An Inspection-Time Analysis

According to mental models theory (MMT) people achieve deduction through three processing stages (Johnson-Laird & Byrne, 1991). In the *comprehension* stage a model representing the information in the premises is constructed. In the *description* stage premises are integrated to formulate a putative conclusion. Finally, in the *validation* stage this conclusion is compared with alternative models of the premises to determine whether it has a counterexample. If no counterexample model is found the conclusion is accepted. Although the MMT of deduction has achieved substantial support from numerous studies (see Johnson-Laird, 2001), disputes remain over whether deduction might be better understood as involving a "mental logic" (e.g., Rips, 1994), or probabilistic Bayesian computations (e.g., Chater & Oaksford, 2001).

The present study tested MMT predictions about premise representation during deduction by monitoring the time people spent processing the components of categorical syllogisms. Such syllogisms consist of two premises and a conclusion, each containing one of four logical quantifiers that determine the problem's *mood* (quantifiers are denoted by letters: A = *All*, E = *No*, I = *Some*, and O = *Some…are not*). A syllogism's validity depends on the mood of the premises and conclusion and the order of the terms in the argument (its *figure*). We follow the convention (Johnson-Laird & Bara, 1984) of referring to syllogism terms as A, B and C, where A is linked to B in the first premise, B to C in the second premise, and A to C in the conclusion. Thus, four figures can be defined: AB-BC, BA-CB, AB-CB, and BA-BC.

Figural Effects in Syllogistic Reasoning

Research has shown that a major influence on performance with categorical syllogisms derives from their figure. Figure can bias both *performance accuracy* and *directionality preferences* for term orders in conclusions (Johnson-Laird & Bara, 1984). Figure AB-BC is easiest to reason with, followed by BA-CB; the remaining figures promote increasing difficulty. The directionality biases seen in generated conclusions take the form of a preference for AC conclusion-term orders with AB-BC arguments and CA orders with BA-CB arguments. The other figures reveal no directional preferences. Examining figural effects is the key concern of the present study as they are critical to understanding the nature of human reasoning. Indeed, Dickstein (1978) argued that a theory of syllogistic inference would be incomplete without accounting for figural biases. Importantly, too, figural effects generalise beyond categorical syllogisms, having been observed in conditional inference (Grosset & Barrouillet, 2003), in tasks combining disjunctive, conjunctive and conditional components (Garcìa-Madruga, Moreno, Carriedo, Gutiérrez, & Johnson-Laird, 2001), and in relational reasoning (Johnson-Laird & Bara, 1984).

Hunter (1957) was one of the first theorists to forward an explanation of premise-integration processes, albeit in the relational inference paradigm. Hunter proposed two mental operations to enable premise combination: (1) the *conversion* of terms within a premise, or (2) the mental *reordering* of whole premises. For the AB-BC figure the middle terms fall adjacent to each other; the second premise easily integrates with the representation of the first, without recourse to mental operations. For the BA-CB figure the middle terms are not contiguous, so the formation of an integrated representation requires mental operations and is more difficult. Johnson-Laird and Bara (1984) adopted Hunter's proposals as part of the premise-integration process within the MMT of syllogistic inference. They consider reordering to be the dominant mental operation for the BA-CB figure, which results in its representation in the form CB-BA. The AB-CB and BA-BC figures are the most complex as they require the conversion of a premise, which is argued to be more cognitively demanding and error prone than premise reordering (Bara, Bucciarelli, & Lombardo, 2001; though see Roberts, Newstead, & Griggs, 2001, for evidence that conversion operations may play only a limited role in syllogistic reasoning). MMT explains directionality effects in conclusion generation by proposing that reasoners, having constructed integrated models where middle terms are contiguous, state conclusions based on a "first-in, first-out" (FIFO) principle (Broadbent, 1958).

Stenning (2002), however, is critical of the MMT explanation of figural effects on conclusion order, and suggests that it is insufficient merely to claim a role for the FIFO principle without converging evidence. Moreover, he argues that substantiation of such a mechanism could not, in itself, be taken as definitive support for any particular representational system such as MMT. Indeed, the mental manipulations proposed by Hunter (1957) could be adopted as a component of any representational account of deduction that specifies the integration of premises. Notwithstanding these arguments, it remains that MMT is the only contemporary account of syllogistic reasoning that *explicitly* posits a working memory demand induced by figure[1] that promotes figural biases on performance. Indeed, the concept of working-memory demand is central to the implementation of the *description* stage in the most recent computational model of MMT (Bara et al., 2001).

Some theorists argue that figural biases are a methodological artefact restricted to conclusion-production paradigms (Geurts, 2003; Rips, 1994; Wetherick, 1989), whilst others dispute that cognitive load is induced by figure (see Chater & Oaksford's, 1999, probability heuristics model--PHM). For example, in response to Espino, Santamaría, and Garcìa-Madruga's (2000a) evidence for figural effects based on the activation of end-terms in memory, Oaksford argues: "Their finding does not address the issue of processing difficulty, so it is difficult to see how it bears on PHM's prediction of no differences in processing difficulty between figures" (Oaksford, 2001, p. 208). Oaksford disputes the importance of Espino et al.'s results for the PHM, but indicates that if a measure of processing difficulty demonstrated a figural demand, this would challenge the current formulation of this model.

Espino, Santamaría, and Garcìa-Madruga (2000b) produced evidence for figure-induced processing demands with BA-CB relative to AB-BC syllogisms. They argued that processing difficulty with figure BA-CB should promote increased inspection times for the *second premise* from a MMT perspective. This prediction was supported using a sequential premise-presentation method that enabled the measurement of processing times for second premises. However, participants were unable to inspect: (1) the first premise after viewing the second, or (2) either premise after the conclusion. This method thus creates an additional working memory burden that may inhibit deployment of preferred reasoning strategies (Gilhooly, Logie, & Wynn, 2002). We note, too, that studies examining the ease of premise integration in temporal and spatial inference by measuring processing times for individual premises have also typically used sequential—rather than parallel—premise presentation (Carrieras & Santamaría, 1997; Morra, 2001; Vandierendonck & de Vooght, 1996), which may constrain the nature of preferred reasoning strategies (though see Roberts, 2000; Schaeken & Johnson-Laird, 2000). Overall, we believe that concerns with the sequential presentation method point to the need for a replication of Espino et al.'s (2000b) study with a more flexible premise-presentation technique.

Processing Direction in Syllogistic Inference

The figural effects we have outlined with categorical syllogisms dominate in the *conclusion-production paradigm*, where premises are given and conclusions generated (e.g., Johnson-Laird & Bara, 1984). In contrast, figural effects are not well established in the *conclusion-evaluation paradigm*, where the validity of presented conclusion is assessed (e.g., Evans, Handley, Harper, & Johnson-Laird, 1999). MMT has tended to finesse distinctions between production and evaluation paradigms, instead emphasizing the stages of *comprehension, description* and *validation*. There is, however, an emerging theoretical debate as to whether processing in evaluation tasks involves either: (1) "forward reasoning" from premises to conclusion, as espoused in standard MMT and recent variants (e.g., Bucciarelli & Johnson-Laird, 1999; Bara et al., 2001; Quayle & Ball, 2000); or (2) "backward reasoning" whereby the presented conclusion is used to guide construction of a model of the premises (e.g., Evans, Handley, & Harper, 2001; Hardman & Payne, 1995; Klauer, Musch, & Naumer, 2000). The backward-reasoning view is clearly rather different to the standard MMT claim that premise-driven processing underpins syllogistic inference[2].

One response to this debate is the suggestion (Morley, Evans, & Handley, 2004) that whilst backward reasoning prevails in conclusion-evaluation tasks (as presented conclusions can motivate inferential processes), forward reasoning dominates in conclusion-production tasks (as there is no given conclusion to guide processing)[3]. This proposal provides a compelling account of why figural effects on response accuracy appear in conclusion-production tasks, but are weak or absent in evaluation tasks. Despite Morley et al.'s (2004) claimed support for this position, few studies have systematically examined figural effects in evaluation tasks (but see Dickstein, 1978, for an account of figural effects in a multiple-choice task). The main evaluation-task studies have been those by Evans et al. (2001) and Quayle and Ball (2000), which focused primarily on belief bias effects rather than figure per se. It may, therefore, be premature to dismiss either the influence of figural effects in the evaluation paradigm or the possibility that forward reasoning processes may dominate in both evaluation and production paradigms.

Overview of Methodology and Predictions

Our experiment was devised to replicate and extend Espino et al.'s (2000b) study by implementing a more flexible process-tracing method. We employed a technique whereby clicking a mouse cursor on designated screen areas revealed syllogistic components and enabled computer-logging of inspection times. Once the cursor was moved from an active screen area the syllogistic component was immediately re-masked. This approach was inspired by mouse- and eye-tracking studies of the Wason Selection Task (Evans, 1996; Ball, Lucas, Miles, & Gale, 2003) and, in particular, the "mouse-contingent display" technique pioneered by Schroyens, Schaeken, Fias, & d'Ydewalle (1999, 2000) for studying on-line propositional reasoning processes (see also Roberts & Newton, 2001).

These methods are all predicated on the assumption that inspection times provide a direct measure of the *processing effort* devoted to problem components being gazed at. In studying reasoning with categorical syllogisms various predictions can be examined with inspection-time data. According to the standard MMT, AB-BC syllogisms are easier to process than BA-CB ones, and should be associated with lower premise inspection times. However, backward-reasoning theories predict no inspection-time imbalance across figures, as conclusion-driven strategies will not entail re-representation of premises in the BA-CB figure to form an integrated model. In addition, standard MMT also predicts that participants prefer AC conclusions for AB-BC syllogisms and CA conclusions for BA-CB syllogisms. Accordingly, non-preferred conclusions may be associated with longer inspection times. Again, no such inspection time imbalance is predicted by theories emphasising backward-reasoning.

In relation to the *validity status* of conclusions it is also possible to derive inspection-time predictions from MMT variants. Some theorists posit no inspection-time differences across components of valid and invalid problems because only a *single* possible model is considered in either case (e.g., Evans et al., 2001). Others (e.g., Quayle & Ball, 2000) predict lower inspection times for components of valid problems because they require the consideration of only a single model to be accepted, whereas invalid problems require further model construction to refute the fallacy. Thompson, Striemer, Reikoff, Gunther, and Campbell (2003) likewise predict lower inspection times for valid conclusions. They suggest (cf. Evans et al., 2001; Klauer et al. 2000) that reasoners only search for a single model that is consistent with a conclusion, but find this easier for valid conclusions (because all models of the premises are consistent with these) than for invalid conclusions (where only some models are consistent). Yet another construal, based on the standard MMT, is that in order to accept a conclusion as valid all possible models require consideration, whereas invalid conclusions require discovery of only a single disconfirming model to be rejected. Thus, interesting oppositional predictions are possible concerning the level of scrutiny that will be directed at the components of valid and invalid problems.

Method

Participants

Forty-four psychology students (aged 18 to 45) from the University of Derby gained course credit for participation. None had prior knowledge of reasoning research.

Design and Materials

The experiment involved a 2 x 2 x 2 repeated-measures design that manipulated Figure (AB-BC vs. BA-CB), Mood (IEO vs. EIO), and Validity (valid vs. invalid). The dependent measures were conclusion-acceptance rates and inspection times for the components of presented syllogisms (i.e., Premise 1, Premise 2, and Conclusion). Participants were presented with eight target syllogisms (four in figure AB-BC; four in figure BA-CB) in the moods IEO and EIO. Half the problems were presented with a logically valid conclusion, and half with an invalid *possible strong* conclusion (Evans et al., 1999). IEO problems have a valid *Some A are not C* conclusion and an invalid *Some C are not A* conclusion, whilst the reverse holds for the EIO problems. There were four practice items in the moods AEA, III, IAI and AEE. The content of all syllogisms was neutral, and involved arbitrary combinations of professions, pastimes, and hobbies. The order of problems was partially counterbalanced and content was systematically rotated through the different problem forms. Authorware 5.1 running on Windows PCs was used to present instructions and problems and to record responses and inspection times.

Procedure

Participants were presented with the following computer-based instructions: "This is an experiment to test people's reasoning ability. You will be shown 16 reasoning problems. For each problem there will be three masked statements. These will be labelled 'Premise 1', 'Premise 2' and 'Conclusion'. By clicking your mouse on the masked areas you can reveal the statements. For each problem you are asked if the conclusion given below the premises may be logically deduced from them. You should answer this question on the assumption that the two premises are, in fact, true. You may revisit each of the three masked areas as many times as you wish although you cannot view more than one area simultaneously. [An example problem was presented at this point]. If, and only if, you judge that the conclusion necessarily follows from the premises, you should click 'Yes'; if you judge that it does not necessarily follow you should click 'No'. Please take your time and be sure that you have the right answer before giving your response. After each trial a box will appear saying 'Click to continue’. Do this when you are ready to proceed. The first four trials are practice items”.

Results and Discussion

Conclusion-Acceptance Rates

Statistical assessment of conclusion-acceptance rates (Table 1) using Analysis of Variance (ANOVA) revealed a reliable main effect of Validity, with valid conclusions being accepted more frequently than invalid ones (68% vs. 52%), $F(1, 43) = 5.09$, $MSE = .541$, $p = .029$. No other effects or interactions were evident. Acceptance rates for the present syllogisms are similar to previous conclusion-evaluation studies (e.g., Evans et al., 1999). The main effect of Validity supports the view that people have a modicum of deductive competence on this task whilst simultaneously being biased toward accepting conclusions not strictly warranted by the given premises (Evans & Over, 1997).

(Table 1 about here)

Inspection Times

Table 2 summarizes inspection-time data relating to syllogistic components collapsed over correct and incorrect conclusion-evaluation responses (cf. Thompson et al., 1993, for evidence that time-based measures in syllogistic inference are comparable between collapsed data and correct-only data). Descriptive analysis revealed that data were positively skewed. A square-root transformation reduced this skew successfully, and statistical analyses were performed on transformed data. For clarity of interpretation we report inspection-time means throughout the paper both before transformation and converted back into original units (seconds) after transformation.

(Table 2 about here)

Premise inspection times. An ANOVA was conducted on premise inspection-time scores using the factors of Figure (AB-BC vs. BA-CB), Validity (valid vs. invalid), and Premise Component (Premise 1 vs. Premise 2), and collapsing across Mood. This revealed a reliable main effect of Figure, $F(1, 43) = 6.08$, $MSE = 0.729$, $p = .018$, with premises in the BA-CB figure inspected for longer than premises in the AB-BC figure. This result indicates increased processing demands associated with premises when middle terms are non-contiguous, and supports the critical MMT prediction that additional mental manipulations are required to align such premise terms (Johnson-Laird & Bara, 1984). The analysis failed to support a main effect of Validity, $F(1, 43) = 2.39$, $MSE = 0.507$, $p = .13$, although the mean cumulative premise inspection time for valid syllogisms (13.97 s after transformation) was less than that for invalid syllogisms (15.23 s after transformation). There was no significant effect of Premise Component, $F(1, 43) = 1.27$, $MSE = .359$, $p = .28$.

An assessment of the specific locus of premise-based inspection-time effects was possible by examining the data for the presence of an interaction between Premise Component (Premise 1 vs. Premise 2) and Figure (AB-BC vs. BA-CB). For syllogisms in the BA-CB figure the greatest inspection time should arise for Premise 1 as this needs to be both refreshed and integrated with the representation of Premise 2 (Johnson-Laird & Bara, 1984), whereas inspection times for two premises in the AB-BC figure should be similar to one another as Premise 2 simply needs to be integrated with the current representation of Premise 1. The mean values depicted in Table 2 suggest that this pattern of premise-inspection times is evident, and an ANOVA revealed that the interaction between Premise Component and Figure was reliable, $F(1, 43) = 4.36$, $MSE = 0.149$, $p = .043$. LSD main effects analyses demonstrated a significant difference between Premise 1 viewing times in the AB-BC figure compared with the BA-CB figure ($p = .007$), with the latter being higher than the former. There was no significant difference between Premise 1 and Premise 2 inspection times for the AB-BC figure, although there was a near significant difference between Premise 1 and Premise 2 inspection times for the BA-CB figure (p = .072), with Premise 1 having the larger mean viewing time, as per MMT predictions.

Conclusion inspection times. Our analyses of conclusion inspection times were devised: (1) to test the MMT prediction of increased inspection times for conclusions in a non-preferred direction relative to a preferred direction (reflecting a conflict between the directionality of presented conclusion terms and the directionality of terms in mental models), and (2) to examine validity effects on conclusion inspection times as the presence of processing differences for valid and invalid conclusions is informative about the status of different contemporary variants of MMT. Conclusion inspection-time data are presented in Table 3, collapsed across Mood and Figure. An ANOVA revealed a reliable main effect of Conclusion Direction, $F(1, 43) = 7.40$, $MSE = .35$, $p = .009$, with non-preferred term directions being inspected for longer than preferred ones. There was also a reliable main effect of Validity, $F(1, 43) = 6.02$, $MSE = .516$, $p = .018$, with invalid conclusions being viewed for *longer* than valid ones. Conclusion Direction and Validity were not found to interact, $F(1, 43) = 0.26$, $MSE = .49$, $p = .61$.

(Table 3 about here)

General Discussion

Our study tested predictions based on the standard MMT of syllogistic reasoning (Johnson-Laird & Byrne, 1991), including the claimed role of forward-reasoning processes that begin with premise modelling and progress to conclusion examination. Under this view, problem figure impacts directly on the processing effort required to formulate models of premises, and figural influences should generalise across conclusion-production and conclusion-evaluation tasks. To date, however, studies that have used an evaluation paradigm (Evans et al., 2001; Morley et al., 2004; Quayle & Ball, 2000) indicate that figure has a limited influence on conclusion acceptance rates, though few studies have examined more subtle *processing-time* measures. Recently, too, it has been suggested that evaluation tasks may be less biased by figure than production tasks, as the former may be dominated by conclusion-driven processes (e.g., Morley et al., 2004).

Our experiment focused on the evaluation paradigm and assessed whether the figure of categorical syllogisms impacts the processing of problems components (i.e., premises and conclusions). The study involved an inspection-time methodology similar to that pioneered by Espino et al. (2000b), but enhanced to allow recursive inspection of previous problem components post-activation to enable detection of the precise locus of figural effects on processing. Moreover, the flexibility to visit and revisit premises and conclusions facilitated the examination of the specific syllogistic components that are associated with relatively longer or shorter inspection-times and the testing of processing effects relating to preferred versus non-preferred conclusion directions predicted under MMT (e.g., Johnson-Laird & Bara, 1984).

A key prediction derived from the standard, premise-driven MMT was that there should be increased viewing of the premises of BA-CB problems relative to AB-BC ones, because BA-CB premises require an additional mental operation (re-ordering Premise 1 after Premise 2) to bring middle terms into contiguity. This finding was supported. MMT also permitted a more fine-grained prediction to be made: For syllogisms in the BA-CB figure, increased viewing times should arise for Premise 1 relative to Premise 2 since it is the representation of Premise 1 that must be refreshed prior to integration with the representation of Premise 2. For the AB-BC figure there should be no marked difference in viewing times for the two premises, which can be integrated in a straightforward manner. The predicted interaction pattern between figure and premise component was evident in the inspection-time data.

A final prediction from MMT related to directionality biases associated with the processing of presented conclusions relative to the assumed order of premise terms in models. We predicted increased inspection times for conclusions whose terms were in a non-preferred direction versus conclusions whose terms were in a preferred direction. This is because participants would be expected to linger longer over conclusions in non-preferred directions, whilst engaging in a process of scanning models in a cognitively demanding right-to-left direction rather than the more natural left-to-right direction assumed under a FIFO principle (Johnson-Laird & Bara, 1984). This effect was reliable. Overall, the MMT view of syllogistic inference as a forward, premise-to-conclusion process achieves good support from our inspection-time data.

We next consider evidence for predictions derived from recent variants of the MMT of syllogistic inference (Hardman & Payne, 1995; Quayle & Ball, 2000), which propose that valid conclusions can be endorsed as necessary without the need for alternative models to be formulated, whereas invalid but “possible” conclusions require multiple-model construction (see Thompson et al., 2003, for an identical prediction based on an alternative MMT account). As such, invalid problems might be expected to be associated with greater premise refreshing and increased conclusion scrutiny in order for models to be actively maintained and fleshed out in working memory, and compared with given conclusions. The predicted validity effect on premise inspection times was in the correct direction but was not significant, whereas the influence of validity on conclusion inspection times was highly reliable.

Overall, our experiment supports key processing assumptions of MMT in its standard form (Johnson-Laird & Byrne, 1991) and some its recent variants that also propose forward-reasoning processes in the evaluation paradigm (e.g., Quayle & Ball, 2000). The figure of a syllogism has detectable influences on premise and conclusion processing in a syllogistic-evaluation paradigm, even though such effects have limited impact upon acceptance rates, which remain stable across the relatively "easy" figures studied here (AB-BC and BA-CB). The validity status of syllogistic conclusions also has a demonstrable effect on conclusion processing and conclusion-acceptance rates.

Evidence for figural influences in an evaluation paradigm runs counter to claims (Morley et al., 2004) for a dissociation between processing direction in conclusion-evaluation versus conclusion-production tasks. Morley et al. propose that forward reasoning dominates in the conclusion-production paradigm, whereas backward reasoning prevails in conclusion-evaluation. Our evidence for figural effects on processing in the evaluation paradigm is not easily reconcilable with this position. However, Morley et al.'s use of belief-oriented materials might explain the discrepant findings across studies (see also Ball, Phillips, Wade, & Quayle, 2006, for eye-tracking evidence of *conclusion-driven* processing in a belief-bias paradigm). Belief-oriented conclusions may compel reasoners toward a conclusion-driven process, whereas neutral conclusions may instead promote a premise-driven process.

Although our evidence for a figure-induced working-memory load is inconsistent with several contemporary theories, we agree with Stenning (2002) that this does not provide *fundamental* evidence against the core principles of accounts that emphasise either rule-based or heuristic processes operating on syntactic aspects of presented information. In fact, the reordering processes proposed by Hunter (1957) could be appended to alternative representational accounts such as those based on mental logic or Euler circles. Moreover, Stenning and Yule (1997) suggest that spatial arrays or verbal lists are plausible explanations of the memory implementation in syllogistic reasoning, although these mechanisms require more detailed specification to account for all of our premise inspection data. Ford (1994) proposed an algebraic substitution procedure that would certainly incur a demand on working memory. However, this explanation was not corroborated by data presented by Espino et al. (2000a) based on their experiments testing the activation of end terms, which supported a premise-reordering manipulation. Finally, some theorists propose that figural bias in syllogistic reasoning stems from a *rhetorical principle* based on the tendency for people to use as the subject of the conclusion a term that has appeared as the subject of one of the premises(e.g., Wetherick & Gilhooly, 1990; see also Chater & Oaksford's, 1999, probability heuristics model). While such theoretical positions have not previously been associated with the notion of a cognitive load, it has recently been suggested (Oberauer, Hörnig, Weidenfeld, & Wilhelm, 2005) that an inadvertent selection of a middle term from the first premise as the subject of the conclusion may occur in the figure BA-CB, and that this putative subject term would require revision upon the reading of the second premise and the realisation that it is, in fact, a middle term. Such a process could explain increases in premise inspection times.

We acknowledge that most theories of syllogistic reasoning might well be able to extrapolate their central principles to provide an account of figural influences on premise and conclusion processing in a way that could accommodate inspection-time effects as seen in the present study. It remains for the advocates of these theories to meet this challenge. However, we would argue that it is to the credit of mental models theory that it alone successfully predicted the full range of inspection-time evidence observed in our study.

References

Ball, L. J., Lucas, E. J., Miles, J. N. V., & Gale, A. G. (2003). Inspection times and the selection task: What do eye-movements reveal about relevance effects? *Quarterly Journal of Experimental Psychology, 56A,* 1053-1077.

Ball, L.J., Phillips, P., Wade, C.N., & Quayle, J.D. (2006). Effects of belief and logic on syllogistic reasoning: Eye-movement evidence for selective processing models. *Experimental Psychology, 53,* 77-86.

Bara B.G., Bucciarelli M., & Lombardo V. (2001). Mental model theory of deduction: A unified computational approach. *Cognitive Science, 25,* 839-901.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon.

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science, 23,* 247-303.

Carreiras, M., & Santamaría, C. (1997). Reasoning about relations: Spatial and nonspatial problems. *Thinking and Reasoning, 3,* 191-208.

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38,* 191-258.

Chater, N., & Oaksford, M. (2001). Human rationality and the psychology of reasoning: Where do we go from here? *British Journal of Psychology, 92,* 193-216.

Dickstein, L. S. (1978). The effect of figure on syllogistic reasoning. *Memory & Cognition, 6,* 76-83.

Espino, O., Santamaría, C., & Garcìa-Madruga, J. A. (2000a). Activation of end-terms in syllogistic reasoning. *Thinking & Reasoning, 6,* 67-89.

Espino, O., Santamaría, C., & Garcìa-Madruga, J. A. (2000b). Figure and difficulty in syllogistic reasoning *Current Psychology of Cognition 19,* 417-428.

Evans, J. St. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. *British Journal of Psychology, 87,* 223-240.

Evans, J. St. B. T., Handley, S. J., & Harper, C. N. J. (2001). Necessity, possibility and belief: A study of syllogistic reasoning. *Quarterly Journal of Experimental Psychology, 54A,* 935-958.

Evans, J. St. B. T., Handley, S. J., Harper, C. N. J., & Johnson-Laird, P. N. (1999). Reasoning about necessity and possibility: A test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1495-1513.

Evans, J. St. B. T., & Over, D. E. (1997). Rationality in reasoning: The problem of deductive competence. *Current Psychology of Cognition, 16,* 3-38.

Ford, M. (1994). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, 54,* 1-71.

Garcìa-Madruga, J. A., Moreno, S., Carriedo, N., Gutiérrez, F., & Johnson-Laird, P. N. (2001). Are conjunctive inferences easier than disjunctive inferences? A comparison of rules and models. *Quarterly Journal of Experimental Psychology, 54A,* 613-632.

Geurts, B. (2003). Reasoning with quantifiers. *Cognition, 86*, 223-251.

Gilhooly, K. J., Logie, R. H., & Wynn, V. (2002). Syllogistic reasoning tasks and working memory: Evidence from sequential presentation of premises. *Current Psychology, 21,* 111-120.

Grosset, N., & Barrouillet, P. (2003). On the nature of mental models of conditional: The case of *If, If then,* and *Only if. Thinking & Reasoning, 9,* 289-306.

Hardman, D. K., & Payne, S. J. (1995). Problem difficulty and response format in syllogistic reasoning. *Quarterly Journal of Experimental Psychology, 48A,* 945-975.

Hunter, I. M. L. (1957). The solving of three term series problems. *British Journal of Psychology, 48,* 286-298.

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, 5,* 434-442.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition 16,* 1-61.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction.* Hove: Lawrence Erlbaum Associates.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107, 852*-884.

Morley, N. J., Evans, J. St. B. T., & Handley, S. J. (2004). Belief bias and figural bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology, 57A,* 666-692.

Morra, S. (2000). On the information-processing demands of spatial reasoning. *Thinking and Reasoning, 7,* 347-365.

Oaksford, M. (2001). Language processing, activation and reasoning: A reply to Espino, Santamaría, and Garcìa-Madruga (2000). *Thinking & Reasoning, 7,* 205-208.

Oberauer, K., Hörnig, R., Weidenfeld, A., & Wilhelm, O. (2005). Effects of directionality in deductive reasoning: II. Premise integration and conclusion evaluation. *Quarterly Journal of Experimental Psychology, 58A*, 1225-1247.

Quayle, J. D., & Ball, L. J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology, 53A,* 1202-1223.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge, MA: MIT Press.

Roberts, M. J. (2000). Strategies in relational inference. *Thinking and Reasoning, 6*, 1-26.

Roberts, M. J., Newstead, S. E., & Griggs, R. A. (2001). Quantifier interpretation and syllogistic reasoning. *Thinking and Reasoning, 7*, 173-204.

Roberts, M. J., & Newton, E. J. (2001). Inspection times, the change task, and the rapid response selection task. *Quarterly Journal of Experimental Psychology, 54A,* 1031-1048.

Schaeken, W., & Johnson-Laird, P .N. (2000). Strategies in temporal reasoning. *Thinking and Reasoning, 6,* 193-219.

Schroyens, W., Schaeken, W., Fias, W., & d'Ydewalle, G. (1999). Is there any need for eye-movement recordings during reasoning? In W. Becker, H. Deubel, & T. Mergner (Eds.). *Current oculomotor research: Physiological and psychological aspects* (pp. 279-285). New York: Plenum Press.

Schroyens, W., Schaeken, W., Fias, W., & d'Ydewalle, G. (2000). Heuristic and analytic processes in propositional reasoning with negatives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1713-1734.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Stenning, K. (2002) *Seeing reason: Image and language in learning to think.* Oxford: Oxford University Press.

Stenning, K., & Yule, P. (1997). Image and language in human information processing: A syllogistic illustration. *Cognitive Psychology, 34,* 109-159.

Thompson, V. A., Striemer, C. L., Reikoff, R., Gunther, R. W., & Campbell, J. D. (2003). Syllogistic reasoning time: Disconfirmation disconfirmed. *Psychonomic Bulletin and Review, 10,* 184-189.

Vandierendonck, A., & De Vooght, G. (1996). Evidence for mental-model-based reasoning: A comparison of reasoning with time and space concepts. *Thinking and Reasoning, 2,* 249-272.

Wetherick, N. E. (1989). Psychology and syllogistic reasoning. *Philosophical Psychology, 2,* 111-125.

Wetherick, N., & Gilhooly, K. (1990). Syllogistic reasoning: Effects of premise order. In K. Gilhooly, M. T. G. Keane, R. Logie, & G. Erdos (Eds.), *Lines of thought: Reflections on the psychology of thinking, Vol. 1.* London: Wiley.

Author Note

We are grateful to Karl Christoph Klauer, Jonathan Evans, Walter Schroyens, Nicki Morley, and Kevin Paterson for helpful comments on earlier versions of this paper. We acknowledge the University of Derby for its financial support.

Footnotes

[1] Ford (1995) suggests an explanation of figural effects in the form of a "verbal algebra". Although this theory does not directly specify working memory involvement, it does lend itself to an implicit assumption that working memory loads may be associated with figural influences on performance.

[2] The contrast between forward and backward processing described here refers to premise-driven versus conclusion-driven processing as defined by Morley, Evans, and Handley (2004), rather than the distinction applied by Dickstein (1978) when contrasting subject-to-predicate and predicate-to-subject processing with classical syllogisms.

[3] We are mindful to refer to the possibility that forward and backward reasoning may *dominate* processing in syllogistic reasoning rather than proposing that reasoning would be observed to be *deterministically* in either form. We recognise the inherent flexibility of much cognitive processing (e.g., Siegler, 1996) such that it would make little sense to predict that all participants will perform the same way on all presented tasks, or, indeed, that individuals will behave the same on all trials. Instead, our research on figural effects focuses on the *prevailing* processing approach used in the evaluation paradigm.

Table 1

Percentage of Conclusion Acceptances as a Function of Figure, Mood, and Conclusion Validity

	Figure AB-BC			Figure BA-CB		
Conclusion validity	IEO	EIO	*M*	IEO	EIO	*M*
Valid	68	55	62	75	71	73
Invalid	54	50	52	45	59	52
M	61	53	57	60	65	63

Table 2

Mean Inspection Times (in Seconds) for Components of Syllogisms as a Function of Figure, Mood, and Conclusion Validity

	Figure AB-BC						Figure BA-CB					
	IEO		EIO		*M*		IEO		EIO		*M*	
Syllogism component	ND	TD	ND	TD	ND	TD	ND	TD	ND	TD	ND	TD
Valid												
Premise 1	6.44	5.20	8.97	7.29	7.70	6.25	7.74	6.97	9.06	7.78	8.40	7.38
Premise 2	6.43	5.62	8.62	6.97	7.52	6.30	8.02	7.18	6.68	5.76	7.35	6.47
Conclusion	5.68	4.93	6.76	5.90	6.22	5.42	5.25	4.67	4.82	4.33	5.03	4.50
Invalid												
Premise 1	7.95	6.50	6.97	6.15	7.46	6.33	10.59	8.82	9.49	8.12	10.04	8.47
Premise 2	8.28	7.13	6.80	5.66	7.54	6.40	8.8	7.78	8.49	7.56	8.64	7.67
Conclusion	7.26	6.10	5.15	4.49	6.21	5.30	6.97	6.20	7.64	6.50	7.31	6.35

Note. ND = natural data in seconds. TD = transformed data (square root of natural data) converted into original measurement units (seconds). Standard errors for the natural data ranged from 0.46 to 1.52, and from 0.11 to 0.20 for transformed data. Standard errors have been omitted from the table to aid readability.

Table 3

Mean Inspection Times (in Seconds) for Conclusions as a Function of Validity and Conclusion Order, Collapsed across Mood and Figure

	Conclusion order					
	Preferred		Non-preferred		*M*	
Validity	ND	TD	ND	TD	ND	TD
Valid	5.25	4.63	6.01	5.23	5.63	4.93
Invalid	6.06	5.35	7.45	6.30	6.76	5.83
M	5.66	4.99	6.73	5.77		

Note. ND = natural data in seconds. TD = transformed data (square root of natural data) converted into original measurement units (seconds).

Standard errors for the natural data ranged from 0.52 to 0.83. For transformed data they ranged from 0.10 to 0.14.