Title
Effects of Visual and Phonological Distinctiveness on Syllogistic Reasoning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN
1069-7977

Authors
Quayle, Jeremy D.
Ball, Linden J.

Publication Date
2007

Peer reviewed

Effects of Visual and Phonological Distinctiveness on Syllogistic Reasoning

Jeremy D. Quayle (j.quayle@shu.ac.uk)
Brain, Behaviour & Cognition Research Group,
Psychology Division, Sheffield Hallam University,
Sheffield, UK

Linden J. Ball (l.ball@lancaster.ac.uk)
Department of Psychology, Lancaster University
Lancaster, UK

Abstract

Mental models theorists (e.g., Johnson-Laird & Byrne, 1991) suggest that syllogism terms are represented in working memory as 'abstract tokens'. However, the role that working memory sub-systems and prior knowledge play in the representation and processing of such terms is poorly specified. Two experiments are reported in which the representational distinctiveness of syllogism terms was manipulated. In Experiment 1 participants were required to evaluate the logical validity of conclusions for syllogisms whose premises contained *visualisable* terms (e.g., spotty or hairy), *character* terms (e.g., friendly or stupid), or *nonsense* terms (e.g., drenful or furplish). A logic x content interaction was observed, such that the effect of logic was greatest with syllogisms whose premises were visualisable and smallest with syllogisms whose premises contained nonsense terms. In Experiment 2 participants were required to evaluate conclusions for syllogisms containing either phonologically similar terms (e.g., fuds, fods and fids) or dissimilar terms (e.g., harks, paps and fids). Again, a logic x content interaction was observed such that a strong effect of logic arose with the dissimilar content, but not with the similar content. Set within a mental models framework, hypotheses are proposed to explain the importance of phonological and visuo-spatial distinctiveness in syllogistic reasoning.

Keywords: Syllogistic reasoning; working memory; distinctiveness.

Introduction

The traditional, categorical syllogism is a deductive reasoning problem comprising two premises and a conclusion. For example:

Some artists are beekeepers
No beekeepers are carpenters
Therefore, Some artists are not carpenters

Within the premises there are three terms: (1) the 'A term' is in the first premise ('artists' in this example), (2) the 'C term' is in the second premise (i.e., 'carpenters'), (3) and the 'B term' is in both premises (i.e., 'beekeepers'). A logically valid conclusion is one that describes the relationship between the A term and C term (referred to as 'end terms') in a way that is necessarily true, given that the premises are true. It is valid as a function of the form or structure of the syllogism, and not because of the content–the words artists, beekeepers and carpenters could be replaced by any nouns or, alternatively, adjectives describing the properties or characteristics of certain classes of entity.

As the two terms in each of the premises can be presented in two possible orders, the terms for a syllogism can be presented in four different arrangements (or 'figures'): A-B, B-C and B-A, C-B (called 'asymmetrical' figures), and A-B, C-B and B-A, B-C (called 'symmetrical' figures). The term 'mood' is used to refer to the different combinations of quantifiers contained within the premises and conclusion. Four different quantifiers are used in standard syllogisms. These are commonly referred to by letters of the alphabet: A = *all*, E = *no*, I = *some*, and O = *some….are not*. The syllogism in the above example, therefore, can be said have the A-B, B-C figure, and the IEO mood.

Since four different quantifiers can be used and there are four different figures, it is possible to produce 64 different premise pairs. However, only 27 of these actually yield a logically valid conclusion (called 'determinate' syllogisms). Whilst people have little difficulty with a few of these determinate syllogisms, many syllogisms are very difficult, and often lead to logically invalid responding from participants. Explaining the systematic patterns of logical and non-logical responding that emerge has been a challenge for theorists in this area.

The Mental Models Theory of Syllogistic Reasoning

The mental models theory of syllogistic reasoning (e.g., Johnson-Laird & Byrne, 1991) provides an account of the patterns of responding that emerge from different forms of syllogism. Its account of deductive reasoning competence and biases has received considerable support in the reasoning literature (e.g., see Evans, Newstead, & Byrne, 1993, for a review). The theory states that individuals begin reasoning with syllogisms by constructing an initial mental model of the two premises. The three terms are represented in the model by 'abstract tokens'. An initial mental model for the example syllogism (above) is shown as follows, using mental models notation:

```
a   [b]
a   [b]
         [c]
         [c]
```

In this notation, arbitrary numbers of letter-tokens are used to represent members of the categories referred to by the three terms. Tokens on the same row share category membership. Hence, this model shows two members of the A term category that are also members of the B term category, and two members of the C term category that are *not* members of the B term category. The brackets around the tokens signify exhaustive representation (i.e., there is no necessity to add further tokens to the model representing these categories). Notice that the A term is not represented exhaustively, suggesting that members of the A term category could exist on different rows of the model. From this initial model individuals generate a conclusion that describes the categorical relationship between the two end-terms. The model above supports the conclusions 'Some A are not C', 'No A are C', 'No C are A' and 'Some C are not A'. The truth of a putative conclusion is then tested against fleshed out versions of the initial mental model (if it is necessary/possible to flesh out the initial model):

```
a   [b]            a   [b]
a   [b]            a   [b]
a        [c]       a        [c]
         [c]       a        [c]
```

An extra token representing the A term has been added to the model on the left to show a situation where some As are Cs. This model falsifies the conclusions 'No A are C' and 'No C are A'. In the model on the right a further A term token as been added to show a possible situation where all Cs are As, thus falsifying the conclusion 'Some C are not A'. If the conclusion cannot be falsified by a fleshed out mental model (e.g., 'Some A are not C'), then it is necessarily true, and is therefore, generated, otherwise it is rejected and a new putative conclusion is generated and tested. A similar process takes place when conclusions are presented for evaluation; if the conclusion cannot be falsified it is accepted, otherwise it is rejected.

Mental models that require fleshing out (i.e., for 'multiple-model' syllogisms) place greater loads on limited working memory than those which do not require fleshing out (i.e., for 'one-model' syllogisms). Reasoning ability is, therefore, constrained by working memory capacity, such that participants more frequently generate valid conclusions to one-model syllogisms than to multiple-model syllogisms (Johnson-Laird & Byrne, 1991). The mental models theory also provides an account of the difficulty associated with certain figures, and how figure affects the forms of conclusion that are generated (see Stupple & Ball, in press).

Working Memory and Syllogistic Reasoning

Whilst Johnson-Laird and colleagues evoke the notion of limited working memory capacity to explain aspects of syllogistic reasoning performance, they do not subscribe to any particular conceptualisation of working memory. Which working memory sub-systems are involved is not stated, although Johnson-Laird has suggested that "…the ability to construct alternative models…should correlate with spatial ability" (Johnson-Laird, 1985, p.190) and that mental images may be a sub-class of mental models (Johnson-Laird, 1983). The suggestion that the generation and evaluation of conclusions involves 'first in first out' scanning of mental models (Johnson-Laird & Byrne, 1991) also implies that models are constructed and manipulated in a visuo-spatial working memory sub-system.

Several studies have investigated the role of working memory sub-systems in deduction (e.g., Duyck, Vandierendonck, & De Vooght, 2003 ; Gilhooly, Logie, Wetherick, & Wynn, 1993; Gilhooly, Logie, & Wynn, 2002; Klauer, Stegmaier, & Meiser, 1997; Quayle & Ball, 2000; Toms, Morris, & Ward, 1993; Vandierendonck & De Vooght, 1997). These studies have, however, yielded somewhat inconsistent results (see Gilhooly, 2005, for a review). For example, Gilhooly et al.'s (1993) investigation of syllogistic reasoning revealed central executive involvement, limited phonological loop involvement, and no visuo-spatial involvement. In contrast, Quayle and Ball's (2000) study indicated more visuo-spatial than phonological involvement.

Content and Representation in Working Memory

Many well-known studies of syllogistic reasoning have used neologisms or nonsense words as the premise terms, in an attempt to avoid the potential confounds associated with terms about which participants might have prior knowledge (e.g., Johnson-Laird & Bara, 1984; Johnson-Laird & Steedman, 1978). The use of neologisms produce what is sometimes termed 'abstract' content. This leads us to consider how such terms might be represented within working memory sub-systems.

Nonsense terms have a sound when they are read, so their representation may have a phonological component. However, they are not obviously visualisable, and they have no associated concept within long-term memory. One possible problem with such words as terms, therefore, is that they are not particularly distinct from one another—they are all *nonsense* words, with nothing but their sound (phonological distinctiveness) and written appearance (graphemic distinctiveness) making them different. Indeed, the terms are so abstract that the boundaries between the categories denoted by the terms may be somewhat imprecise, making such syllogisms difficult to represent and process within working memory. Consequently, the construction, manipulation and scanning of mental models in order to generate and test the validity of putative conclusions may be more difficult than with syllogisms containing real words as terms.

'Real' terms within syllogisms may be any nouns or adjectives that denote category membership. They may be classes of people, animals or objects (e.g., 'artists', 'cats' and 'vehicles'), or words that describe these classes (e.g., 'brave', 'tall' or 'red'). To varying degrees, these words are phonologically distinct from one another, and are directly or indirectly associated with visualisable concepts—a characteristic that may facilitate the reasoning process, since they more readily lend themselves to mental representation than nonsense terms. For example, adjectives such as ‘tall’, ‘spotty’, ‘green’ and ‘thin’ can have a phonological representation and a visual representation. It could be argued that these characteristics make the terms more distinct from one another than nonsense terms. Such visual distinctiveness would make the representational boundaries between the three categories clearer. Consequently, it would be easier to construct, manipulate and scan a mental model of a syllogism containing such terms than would be the case with a syllogism containing nonsense terms. Syllogisms with visual content would, therefore, result in superior reasoning performance in comparison to syllogisms with nonsense content. For example, in a conclusion evaluation task, where participants are required to evaluate the validity of presented valid and invalid conclusions, syllogisms with visual content would be predicted to yield a greater effect of logic than syllogisms with nonsense content.

Experiment 1

In Experiment 1 we set out to test the prediction that syllogisms with visual content will yield a greater effect of logic than syllogisms with abstract/nonsense content. An experiment in which reasoning performance with these two types of content is compared would, however, be confounded. Visual terms may not only lend themselves more readily to visual representation than nonsense terms, but they are also associated with concepts already existing with long-term memory, whilst nonsense terms are not. The involvement of pre-existing concepts may work, for example, to ease the representational loads placed on limited working memory capacity. Hence, if a difference in performance was found between visual and nonsense syllogisms, then this may not be because of a visual/non-visual distinction, but because the visual terms are pre-existing concepts and nonsense terms are not.

To arbitrate between these two possibilities, participants may also be presented with syllogisms containing abstract adjectives that are non-visual, pre-existing concepts, such as words describing personality characteristics or dispositional traits (e.g., friendly, brave or eccentric). We would argue that these character terms, not being directly associated with visual information, are less distinct from one another than visual terms, and thus will yield a smaller effect of logic than visual terms. Character adjectives do, however, have pre-existing representations in long-term memory, which could mean that they will yield a greater effect of logic than the nonsense adjectives.

Method

Participants

An opportunity sample comprising 33 female and 27 male participants was tested. The mean age of participants was 38.6 years (*sd* = 17.1). None of the participants had taken formal instruction in logic and all were tested individually.

Materials

Four forms of multiple-model syllogism were presented. These were in the asymmetrical A-B, B-C and B-A, C-B figures, and in the IEO and EIO moods. Each form of syllogism was presented with its logically valid conclusion and with an ‘indeterminately invalid’ conclusion (i.e., one that is *consistent* with the premises, but not *necessitated* by them). Hence, conclusions were presented in both the C-A and A-C forms.

The premises of the syllogisms amounted to descriptions of fictitious monsters. By using this type of content it was possible to have visual, nonsense and character syllogisms that were comparable in their essential form. To be able to do this, the A terms were the names of monsters (e.g., grobbles, flurbs and broggs), and the B and C terms were adjectives. The ‘visual’ syllogisms contained B and C terms that referred to visual and spatial concepts (e.g., tall, spotty and red). The ‘nonsense’ syllogisms contained B and C terms that were nonsense words designed to sound like adjectives (e.g., wurtly, drenful and geltric). The ‘character' syllogisms contained B and C terms that referred to traits that the monsters might possess (e.g., brave, polite and dangerous). The four forms of syllogism were each presented with the three types of content, making 12 problems in total.

Design

A repeated-measures design was used, with all participants receiving the 12 syllogisms. These were preceded by three practice, one-model syllogisms. The 12 experimental problems were presented in a random order, which was rotated so that each problem appeared once in each serial position, creating 12 versions of the test booklet. There were two independent variables: (1) *logic* (two levels: valid vs. invalid), and (2) *content* (three levels: nonsense vs. visual vs. character). Participants were required either to accept or reject presented conclusions (the dependent variable).

Procedure

Participants were presented with the syllogisms in printed test booklets. The following instructions were presented on the second page:

“This is an experiment to test people's reasoning ability. You will be given 15 problems. On each page, you will be shown two statements describing monsters (called broggs, grobbles, flurbs and zutters). Some of the characteristics in the statements you will recognise (e.g., tall, spotty, clever), and some you will not (e.g., lurthy, drenful, geltric) as they are unique to these monsters. You are asked if certain conclusions (given below the statements) may be logically deduced from the two statements. You should answer this question on the assumption that the two statements are, in fact, true. If, and only if, you judge that the conclusion *necessarily* follows from the statements, you should tick the ‘true’ box, otherwise the ‘false’ box. Please take your time and be sure that you have the right answer before moving on to the next problem. You must not make notes or draw diagrams to help you in this task”.

Results

The percentages of conclusions accepted as a function of logic and content are presented in Table 1. A Wilcoxon signed-ranks test showed that overall, significantly more valid conclusions were accepted than invalid ones (z = 3.71, p < .001, one-tailed). The effect of logic was also significant for each type of content: *visual* (z = 4.15, p < .001, one-tailed); *character* (z = 3.24, p < .001, one-tailed); and *nonsense* (z = 1.84, p < .05, one-tailed).

Table 1.
Percentages of conclusions accepted as a function of logic and content in Experiment 1.

	Content			Overall
	Visual	Character	Nonsense	
Valid	83	81	75	80
Invalid	53	61	62	59
Difference	30	20	13	21

In order to test for the interaction between logic and content, scores for the invalid problems were subtracted from scores for the valid problems across participants to give an index of the size of the effect of logic for the visual, character and nonsense syllogisms. A Wilcoxon signed-ranks test was used to see if the size of the effect of logic differed significantly between visual and nonsense syllogisms. This showed a significant interaction between logic and content, such that the size of the effect of logic was greater with the visual syllogisms than with the nonsense syllogisms (z = 3.71, p < .05). A Friedman's Chi-square test was then used to test for a logic × content interaction, when the three types of content are considered together. Whilst the character syllogisms yielded a smaller effect of logic than the visual syllogisms but a greater effect of logic than the nonsense syllogisms (see Table 1), this interaction fell short of significance (χ^2 (df = 2), p = .077).

Discussion

Our main prediction was supported, in that the visual problems resulted in a greater effect of logic than the nonsense problems, thus supporting the idea that the visualisability of syllogism terms affects the ability to construct, manipulate and scan mental models of the premises. These results would also seem to provide further support for the involvement of visuo-spatial working memory in syllogistic reasoning (cf. Quayle & Ball, 2000).

When all three forms of content were entered into the analysis, the interaction between logic and content was only marginally significant, although, as can be seen from Table 1, the trend is such that the effect of logic is greater with the visual problems than with the character problems, and greater with the character problems than with the nonsense problems. A possible explanation for this observed pattern in the effect of logic considers how the terms vary in relation to three characteristics: (1) their capacity to have a phonological representation; (2) whether they are associated with pre-existing concepts in long-term memory–that potentially ease loads on working memory; and (3) their capacity to have a visuo-spatial representation. The nonsense terms can be represented in phonological working memory, but only in an abstract fashion in visuo-spatial working memory. These terms do not evoke pre-existing concepts within long-term memory. Hence, nonsense terms produce the smallest effect of logic. The character terms share the phonological and visuo-spatial characteristics of the nonsense terms. However, they are associated with pre-existing concepts, causing them to yield a greater effect of logic than the nonsense terms. Representation of the visual terms can involve phonological and visuo-spatial working memory sub-systems, as well as the evocation of concepts within long-term memory. Hence, they produce the largest effect of logic.

Alternatively, these findings might be explained purely in terms of varying degrees of distinctiveness. That is, it is not that terms vary in the extent to which they can be visualized or relate to pre-existing concepts, but that they vary simply in their distinctiveness. The visual terms are the most distinctive, since they correspond to fairly unambiguous visual and spatial concepts that pre-exist in long-term memory. The character terms are the next most distinctive—whilst they are abstract terms, they will be indirectly associated with concrete concepts. For example, the word 'brave' might be associated with soldiers, medieval knights or specific daring acts. Nonsense terms, however, are the most abstract, having no obvious visual representation and no obvious prior associations. The distinctiveness of terms may affect the ease with which premises can be represented within a mental model, irrespective of their visuo-spatial or phonological qualities, and thus more distinctive terms yield greater effects of logic than less distinctive terms.

Experiment 2

In Experiment 2 we set out to test the idea that differences in the size of the logic effect between the different types of content in Experiment 1 can be explained in terms of distinctiveness alone, rather than visualisability or the easing of working memory demands associated with pre-existing concepts. The visual, character and abstract terms in Experiment 1 were phonologically distinctive (i.e., for each content type, the adjectives in a syllogism were not phonologically similar to each other). The content varied in terms of *visual* distinctiveness, and distinctiveness in the evocation of pre-existing concepts. In Experiment 2, therefore, we aimed to test for the effect of purely *phonological* distinctiveness on syllogistic reasoning, controlling for visual distinctiveness and the possible effect of pre-existing concepts by using only nonsense terms. If the results of Experiment 1 are attributable to distinctiveness, then we would expect syllogisms with phonologically dissimilar content to yield a greater logic effect than those with phonologically similar content.

Method

Participants

An opportunity sample of 41 female and 23 male participants was tested. The mean age of participants was 28.27 years (*sd* = 11.63). None of the participants had taken formal instruction in logic and all were tested individually.

Materials

The logical forms of the problems used in Experiment 2 were identical to those used in Experiment 1. The terms were all one-syllable, nonsense adjectives. In this way, word-length was controlled, and it was simple to produce terms that were either phonologically similar or phonologically dissimilar. The phonologically similar terms within a syllogism were words with the same beginning and end consonants, but different middle vowels (e.g., juks, jeks and jiks). The phonologically dissimilar terms were words with different beginning and end consonants, and also different middle vowel sounds (e.g., zaps, toks, and yugs) The four forms of syllogism were each presented with the two types of content, making eight problems in total.

Design

A repeated-measures design was used, with all participants receiving the eight syllogisms. These were preceded by two practice, one-model syllogisms. The eight experimental problems were presented in a random order, which was rotated so that each problem appeared once in each serial position (creating eight versions of the test booklet). There were two independent variables: (1) *logic* (two levels: valid vs. invalid), and (2) *content* (two levels: similar vs. dissimilar). Again, conclusion acceptance or rejection was the dependent variable.

Procedure

Participants were presented with the syllogisms in printed test booklets. The instructions were the same as those used in Experiment 1, except for the sentence explaining how the syllogisms were descriptions of monsters, which read: "You will be given 10 problems. On each page, you will be shown two statements describing monsters".

Results

The percentages of conclusions accepted as a function of logic and content are presented in Table 2. A Wilcoxon signed-ranks test showed that overall, significantly more valid conclusions were accepted than invalid ones (z = 2.03, p < .05, one-tailed). The effect of logic was significant for the *dissimilar* problems (z = 2.91, p < .01, one-tailed). There was, however, no significant effect of logic with *similar* problems (z = 0.81, p = .21, one-tailed).

Table 2.
Percentages of conclusions accepted as a function of logic and content in Experiment 2.

	Content		
	Similar	Dissimilar	Overall
Valid	72	78	75
Invalid	68	61	65
Difference	4	17	11

To test for the interaction between logic and content, scores for the invalid problems were subtracted from scores for the valid problems across participants to give an index of the size of the effect of logic for the similar and dissimilar syllogisms. A Wilcoxon signed-ranks test was used to see if the size of the effect of logic differed significantly between similar and dissimilar syllogisms. This showed a significant interaction between logic and content, such that the size of the effect of logic was greater with the dissimilar syllogisms than with the nonsense syllogisms (z = 2.09, p < .05).

Discussion

The finding that phonologically dissimilar content yielded a greater effect of logic than phonologically similar content supports a representational distinctiveness explanation for the results of Experiment 1. It would seem that phonologically distinctive terms as well as visuo-spatially distinctive terms are easier to represent within working memory than non-distinctive terms. In the context of the mental models theory of syllogistic reasoning, it may be that distinctiveness affects the ease with which models are constructed, manipulated and scanned, which in turn affects reasoning performance. These results also provide further support for the involvement of phonological working memory in syllogistic reasoning (cf. Gilhooly et al., 1993). Below we offer a detailed explanation for our findings that integrates the theoretical ideas considered so far.

General Discussion

In accounting for syllogistic reasoning performance and biases, the mental models theory invokes assumptions concerning working memory, such as its limited processing capacity and its role in the serial storage and scanning of information. The theory also states that the terms within syllogisms are represented in working memory by 'abstract' tokens. However, the representational sub-system(s) where these abstract tokens are stored are not explicitly specified, although Johnson-Laird has indicated that model construction may take place in some spatial medium. By manipulating the phonological and visuo-spatial nature of terms within syllogisms we have provided further support for the involvement of both phonological and visuo-spatial working memory in syllogistic inference. Our results, therefore, accord with previous findings as reported by Gilhooly et al. (1993) and Quayle and Ball (2000).

Based on our results we would suggest that the notion of 'abstract tokens' referred to in the mental models theory is unsatisfactory, since the tokens representing term categories within mental models are only purely abstract when no distinctive phonological, visual or spatial information is available to the individual. The data from our experiments support the idea that without this information individuals struggle to construct, manipulate and scan mental models. This is because the representational boundaries between categories are vague, so that the processing of tokens becomes a muddled endeavor. Such an explanation can account for the generally poor performance of participants in Experiment 2 when evaluating conclusions to syllogisms containing phonologically, visually and spatially indistinct terms. With these problems participants were mostly unable to establish the validity of presented conclusions.

When phonological, visual or spatial information is available, tokens representing terms become less abstract. This information makes the representational boundaries between categories more distinct, such that the processing of tokens is a less muddled activity. This can explain how the greatest effect of logic was observed with syllogisms containing terms referring to visual or spatial concepts in Experiment 1. With these materials phonological, visual and/or spatial information was available to bolster representations of categories within working memory (see Capon, Handley, & Dennis, 2003, for related evidence from an individual differences study). In this instance, the difference between valid and invalid conclusion acceptances was 30%—over seven times greater than for the phonologically similar terms in Experiment 2.

The second largest effect of logic was observed with syllogisms containing terms referring to character concepts. With these materials phonological information was available, but no obvious visual or spatial information. However, since the terms are linked with concepts stored within long-term memory, they are indirectly associated with visuo-spatial information, which may serve to bolster the mental representation of categories. The third largest effect of logic was observed with syllogisms containing phonologically dissimilar nonsense terms. If we assume that graphemic variations between words are of little help in constructing mental models, then with these materials the only information that obviously distinguishes between term categories is their phonological code.

Finally, the results of these experiments suggest that further insights into the role of working memory in syllogistic reasoning may be gained through exploration of the relative importance of phonological and visuo-spatial distinctiveness in syllogistic reasoning performance, as well as the relative importance of visual and spatial distinctiveness. A similar method to that used here may prove fruitful, although we acknowledge that devising suitable materials that facilitate these comparisons may prove to be quite a challenge.

Acknowledgments

We thank Richard Nakielny, Karen Matthews and Rachael Tidswell for their assistance in the collection of data.

References

Capon, A., Handley, S.J., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking and Reasoning, 9,* 203—244.

Duyck, W., Vandierendonck, A., & De Vooght, G. (2003). Conditional reasoning with a spatial content requires visuo-spatial working memory. *Thinking and Reasoning*, 9, 267—287.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Hove, U.K.: Lawrence Erlbaum Associates Ltd.

Gilhooly, K. J. (2005). Working memory and strategies in reasoning. In M. J. Roberts and E. J. Newton (Eds.), *Methods of thought: Individual differences in reasoning strategies.* Hove: Psychology Press.

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition*, *21*, 115—124.

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

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge: Cambridge University Press.

Johnson-Laird, P. N. (1985). Deductive reasoning ability. In R. J. Sternberg (Ed.), *Human abilities: An information processing approach*. New York: W.H. Freeman.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition, 16,* 1—62.

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

Johnson-Laird, P. N., & Steedman, M. (1978). The psychology of syllogisms. *Cognitive Psychology, 10*, 64—99.

Klauer, K. C., Stegmaier, R., & Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking and Reasoning, 3,* 9—47.

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.

Stupple, E. J. N., & Ball, L. J. (in press). Figural effects in a syllogistic evaluation paradigm: An inspection-time analysis. *Experimental Psychology.*

Toms, M., Morris, N., & Foley, P. (1994). Characteristics of visual interference with visuospatial working memory. *British Journal of Psychology, 85,* 131—144.

Vandierendonck, A., & De Vooght, G. (1997). Working memory constraints on linear reasoning with spatial and temporal contents. *Quarterly Journal of Experimental Psychology, 50A,* 803—520.