Memory & Cognition
2009, 37 (6), 759-768
doi:10.3758/MC.37.6.759

Phonological and visual distinctiveness effects in syllogistic reasoning: Implications for mental models theory

Linden J. Ball
Lancaster University, Lancaster, England

and

Jeremy D. Quayle
Sheffield Hallam University, Sheffield, England

Two experiments are reported in which the representational distinctiveness of terms within categorical syllogisms was manipulated in order to examine the assumption of mental models theory that abstract, spatially based representations underpin deduction. In Experiment 1, participants evaluated conclusion validity for syllogisms containing either phonologically distinctive terms (e.g., *harks*, *paps*, and *fids*) or phonologically nondistinctive terms (e.g., *fuds*, *fods*, and *feds*). Logical performance was enhanced with the distinctive contents, suggesting that the phonological properties of syllogism terms can play an important role in deduction. In Experiment 2, participants received either the phonological materials from Experiment 1 or syllogisms involving distinctive or nondistinctive visual contents. Logical inference was again enhanced for the distinctive contents, whether phonological or visual in nature. Our findings suggest a broad involvement of multimodal information in syllogistic reasoning and question the assumed primacy of abstract, spatially organized representations in deduction, as is claimed by mental models theorists.

The quest to understand people's reasoning with categorical syllogisms has been active for many years (see Evans, Newstead, & Byrne, 1993, for a review), and studies in this area continue unabated (e.g., Espino, Santamaría, Meseguer, & Carreiras, 2005; Geurts, 2003; Oberauer, Hörnig, Weidenfeld, & Wilhelm, 2005). *Categorical syllogisms* are deductive problems 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: the *A term* in the first premise (artists), the *C term* in the second premise (carpenters), and the *B term* in both premises (beekeepers). A valid conclusion describes the relationship between the A and C 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, not because of the content.

The terms within syllogisms can appear in four different arrangements or *figures*: A–B, B–C and B–A, C–B for *asymmetrical* figures, and A–B, C–B and B–A, B–C for *symmetrical* figures. The term *mood* is used to refer to the different combinations of quantifiers within the premises and conclusion. The four quantifiers in standard syllogisms are denoted by letters of the alphabet: A, *all*; E, *no*; I, *some*; and O, *some . . . are not*. The example syllogism above has the A–B, B–C figure, and the IEO mood. Although people have little difficulty with certain syllogisms, many others are difficult and promote nonlogical responses. Explaining the patterns of logical and nonlogical performance that emerge with categorical syllogisms has been a major theoretical challenge, and, in grappling with conceptual issues, theorists have often made assumptions about the mental representations that underpin syllogistic inference. In this article, we examine the representational assumptions of the mental models theory of syllogistic reasoning (e.g., Johnson-Laird & Bara, 1984; Johnson-Laird & Byrne, 1991; see also Bara, Bucciarelli, & Lombardo, 2001, for recent refinements and extensions).

The mental models theory of syllogistic inference continues to dominate the literature, not least because of the considerable support that it has received from experimental studies of both reasoning development (e.g., Bara, Bucciarelli, & Johnson-Laird, 1995) and adult performance (e.g., Bucciarelli & Johnson-Laird, 1999). Furthermore, unlike other theories (e.g., Rips's [1994] rule-based account, and Chater & Oaksford's [1999] probability heuristics model), the mental models theory can provide compelling explanations of two central phenomena associated with categorical syllogisms: the striking impact of figure on premise-processing latencies (e.g., Espino et al., 2005; Stupple & Ball, 2005, 2007) and the systematic influence of conclusion believability on acceptance rates and problem processing times (e.g., Ball, Phillips, Wade, & Quayle, 2006; Garnham & Oakhill, 2005;

L. J. Ball, l.ball@lancaster.ac.uk

Quayle & Ball, 2000; Stupple & Ball, 2008). Other theories fail to show this breadth of explanatory capability.

One further aspect of the mental models theory that makes it particularly amenable to our interest in representational issues in syllogistic reasoning is that it embodies clear assumptions about the representations underpinning deduction. For example, in explaining how people evaluate conclusions to presented premises, the theory assumes that individuals begin by constructing an initial model of the premises, in which the terms and their categorical relations are represented as abstract tokens organized within two-dimensional spatial arrays (Johnson-Laird, 1996, 1998, 2005). Such models, moreover, are not identified with visual images (because they are abstract), although Johnson-Laird (e.g., 1998) suggested that it may be possible for people to construct an image of what a model represents from a certain point of view.

To clarify the nature of mental model representations in reasoning, consider the initial model that an individual might construct for the premises shown earlier. This initial model might take the following form (using Johnson-Laird & Byrne's [1991] 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. The tokens on the same row share category membership. Hence, this model shows two members of the artists category who are also members of the beekeepers category and two members of the carpenters category who are not members of the beekeepers category. The brackets around the tokens signify exhaustive representation (i.e., it is not possible to add further tokens to the model for these categories). Note that the A term is not represented exhaustively, suggesting that members of the artists category could exist on different rows of the model. Having constructed this initial model, the reasoner can then determine whether it supports the presented conclusion (*Some artists are not carpenters*), which it does. The necessity of this conclusion must, however, be tested against fleshed-out versions of the initial mental model (such as the following) to check whether a counterexample model is possible:

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

In the left-hand model, an extra token representing the A term has been added to show a situation in which some artists are carpenters. This model still supports the given conclusion that some artists are not carpenters. In the model on the right, a further A-term token has been added to show a possible situation in which all carpenters are artists, and, again, the given conclusion that some artists are not carpenters holds in this final model. Whenever a conclusion is not falsified by fleshed-out mental models, it is valid; otherwise, it is invalid.

The view that mental models are constructed as abstract tokens within a spatial substrate has recently gained support from various lines of research. For example, evidence for figural biases and conclusion-order preferences in syllogistic inference (Espino, Santamaría, & García-Madruga, 2000; Espino et al., 2005; Stupple & Ball, 2005, 2007) is readily interpretable in terms of extracting information from spatially organized representations. More compelling still is evidence that congenitally blind individuals seem to be able to construct spatially based mental models during reasoning despite their lack of visual experience (e.g., Fleming, Ball, Ormerod, & Collins; 2006; Knauff & May, 2006). Yet another line of evidence comes from studies demonstrating that visual mental imagery invoked by problem contents can actually hinder people's capacity to construct and use the abstract spatial models necessary for effective reasoning—the so-called *visual imagery impedance hypothesis* (e.g., Knauff & Johnson-Laird, 2002; Knauff & May, 2006; Knauff & Schlieder, 2005; see also Bacon, Handley, & McDonald, 2007).

The research of Knauff and colleagues has been particularly valuable in revealing a potential problem with previous studies that have demonstrated inconsistent links between imagery and deduction (e.g., Clement & Falmagne, 1986; De Soto, London, & Handel, 1965; Johnson-Laird, Byrne, & Tabossi, 1989; Shaver, Pierson, & Lang, 1975; Sternberg, 1980). Knauff and Johnson-Laird (2002) suggested that this inconsistency derives from an inherent confounding in such studies between materials invoking *visual* imagery and materials invoking *spatial* representations. Their proposal was that studies revealing enhanced reasoning inadvertently increased the spatial basis of problem contents, whereas studies showing decrements in reasoning (or sometimes no effect) tended to use materials that invoked visual imagery. Knauff and Johnson-Laird's systematic review of previous experiments uncovers persuasive evidence for accepting the viability of their proposal. Likewise, their own empirical evidence for visual impedance in deduction is compelling. Nevertheless, we note that this evidence comes from deductive tasks (i.e., three- and four-term series problems) in which the visual and spatial properties of *relations* have been manipulated (i.e., visuospatial relations such as *above–below* were contrasted with visual relations such as *cleaner–dirtier* and control relations such as *smarter–dumber* that are neither visual nor spatial). It thus remains unclear whether the visual impedance observed with relations will generalize to studies in which it is the visual properties of the actual *terms* within problems that are manipulated.

One aim of the present research was, therefore, to address this latter issue via a content manipulation, whereby syllogism terms were either visually distinctive or visually nondistinctive. It may be the case, for example, that visually distinctive mental tokens are advantageous for deductive reasoning, since these distinctive tokens are not so easily confused within a limited-capacity working memory system (see Miyake & Shah, 1999, for detailed discussion of the working memory concept). There is, in fact, a range of empirical evidence pointing to the benefits of visual distinctiveness in working memory that gives grounds for predicting that such distinctiveness may also be advantageous for the maintenance and manipulation of representations in deductive reasoning. For example, research on immediate recall of unfamiliar Chinese characters shows a *visual similarity effect*, whereby people's recall reveals confusions for characters that are visually similar to each other (Hue & Ericsson, 1988). This effect also arises for immediate recall of visually similar words (e.g., *fly*, *cry*, *dry*) relative to visually distinct words (e.g., *guy*, *sigh*, *lie*), as was demonstrated by Logie, Della Sala, Wynn, and Baddeley (2000). Further evidence for the visual similarity effect comes from developmental research (Hitch, Halliday, Schaafstal, & Schraagen, 1988), in which young children show confusion errors in recognition memory for visually similar pictures (e.g., a pen, a rake, and a brush) relative to visually distinct pictures (e.g., a pen, a ball, and a pig).

In summary, then, the evidence for visual similarity effects in immediate memory retrieval suggests that visually distinctive terms may also be beneficial in maintaining visually based mental representations in deductive reasoning. Moreover, if a visual distinctiveness manipulation were indeed observed to have an advantageous effect on deductive accuracy, the assumption that deduction is always based on models involving highly abstract entities would seem questionable. On the other hand, if the visual impedance hypothesis captures a generic inhibitory effect on model-based reasoning that arises because of visual distraction, this distraction should presumably occur more with visually distinctive terms that lend themselves to imagery-based representations than with visually nondistinctive terms that should be coded using more abstract representations.

A second aim of the present research was to explore the influence of distinctive phonological representations in syllogistic inference. Our interest here parallels that described above in relation to visualizable terms; that is, do phonologically distinctive terms within syllogisms impede or facilitate reasoning? It is possible that phonologically distinctive terms might be beneficial for mental model construction and reasoning, since such distinctiveness would help clarify the nature of category membership denoted by such terms and facilitate the maintenance of information in working memory. In contrast, since the mental models theory emphasizes the role of spatially organized abstract tokens in deduction, the inherent phonological distinctiveness of presented terms might be expected to have a distracting effect on reasoning along similar lines to that proposed according to the visual impedance hypothesis. Again, the working memory literature provides evidence to motivate the prediction that phonological distinctiveness of presented terms may, in fact, be beneficial for reasoning. In particular, a key phenomenon that has long been established in relation to working memory retrieval is the *phonological similarity effect* (e.g., Baddeley, 1966; Conrad, 1964; Conrad & Hull, 1964), whereby immediate serial recall of items that have a similar sound (e.g., the words *cat*, *map*, *man*, *cap*, *mad*) is much more difficult than immediate serial recall of items that have a dissimilar sound (e.g., *pit*, *day*, *pen*, *cow*, *hot*). As with the visual similarity effect, the phonological similarity effect is likewise assumed to arise because similar items have fewer distinguishing features and, hence, are likely to be confused while being maintained within a limited-capacity working memory system (Baddeley, Eysenck, & Anderson, 2009).

EXPERIMENT 1

In Experiment 1, we set out to address the phonological distinctiveness issue described above; that is, do phonologically distinctive syllogistic terms have a facilitatory or distracting effect on deductive reasoning in comparison with phonologically nondistinctive terms? To avoid confounds arising from presented terms being associated either with preexisting concepts in long-term memory or with visualizable objects or entities, all terms in Experiment 1 concerned short nonsense words (e.g., *jeks*, *toks*, *bebs*).

Method

Participants. An opportunity sample of 55 female and 29 male participants was tested. The mean age of the participants was 27.36 years (*SD* = 11.18). None of the participants had taken formal instruction in logic, and all were tested individually.

Materials. Eight multiple-model target syllogisms were presented to each participant, four in the A–B, B–C figure and IEO mood, and four in the B–A, C–B figure and EIO mood. For both figures, half of the conclusions were in the C–A direction and half in the A–C direction. This ensured that half of the conclusions were logically valid and that the other half were indeterminately invalid (i.e., consistent with the premises but not necessitated by them). Across both figures, half of the valid syllogisms involved phonologically distinctive contents, and half involved phonologically nondistinctive contents. The same content manipulation was applied to the invalid syllogisms.

The terms within the syllogisms were all one-syllable nonsense adjectives. In this way, word length was controlled, and it was simple to produce terms that were either phonologically nondistinctive or phonologically distinctive. The phonologically nondistinctive terms were words with the same onset and coda consonants, but with different middle vowels (e.g., *juks*, *jeks*, and *jiks*). The phonologically distinctive terms were words with different onset and coda consonants and also different vowel sounds (e.g., *zaps*, *toks*, and *yugs*). Four sets of phonologically distinctive terms were generated, as well as four sets of phonologically nondistinctive terms (see Table 1). Appendix A lists the full set of eight target syllogisms used in Experiment 1.

To validate the effectiveness of our phonological distinctiveness manipulation, we carried out a pretest using 15 undergraduate students who received payment for their participation. Each participant was given a booklet containing the four sets of distinctive phonological terms and the four sets of nondistinctive phonological terms, shown in Table 1. Each page of the booklet presented a single set of three terms, with a series of rating tasks below the terms. The scales for these rating tasks were 100-mm horizontal lines with labeled endpoints. The participants were asked to register a judgment on

Table 1
Nonsense Words Used As Syllogistic Terms in Experiment 1

Phonological Distinctiveness			
Nondistinctive		Distinctive	
Word Set	Words	Word Set	Words
1	*bubs*, *bebs*, *babs*	5	*zaps*, *toks*, *yugs*
2	*fuds*, *fods*, *feds*	6	*fubs*, *haps*, *beks*
3	*horks*, *herks*, *harks*	7	*paps*, *harps*, *fids*
4	*jeks*, *juks*, *jiks*	8	*yogs*, *keps*, *zuks*

each scale with a vertical line. The order of the three terms on each page was independently randomized for each participant, as was the order of the term sets within each booklet. The participants were asked to imagine that the presented words denoted the names of fictitious monsters.

The first rating task was in response to the question "How phonologically distinctive are the *spoken* forms of these words?" with the presented scale ranging from *not at all phonologically distinctive* to *highly phonologically distinctive*. The scores indicated a strong separation in the expected direction between the phonologically distinctive items ($M = 64.3$, $SD = 21.3$) and the phonologically nondistinctive items ($M = 27.6$, $SD = 18.6$) [$F(1,14) = 23.90$, $MS_e = 421.86$, $p < .001$, $\eta_p^2 = .63$].

The second rating task was in response to the question "To what extent do these words relate to real words that you are familiar with?" with the scale ranging from *not at all related* to *highly related*. The scores here supported our expectation that the participants would view neither the phonologically distinctive item sets ($M = 42.2$, $SD = 16.3$) nor the phonologically nondistinctive item sets ($M = 49.8$, $SD = 18.2$) as relating strongly to familiar words, with there also being no reliable separation in ratings between item sets [$F(1,14) = 2.92$, $MS_e = 149.03$, $p = .11$, $\eta_p^2 = .17$].

The final rating task asked the question "To what extent do these words allow you to build up vivid mental images of the fictitious monsters that they denote?" with the scale ranging from *very easy to build up vivid mental images* to *very difficult to build up vivid mental images*. Again, there was no reliable difference between the visualizability of terms in the phonologically distinctive item sets ($M = 40.6$, $SD = 19.1$) and the visualizability of terms in the phonologically nondistinctive item sets ($M = 43.2$, $SD = 15.6$) [$F(1,14) = 0.14$, $MS_e = 374.01$, $p = .71$, $\eta_p^2 = .01$], with the scores indicating that relatively low visualizability of terms was the norm.

Overall, the pretest data support the view that our two sets of terms were effectively differentiated in relation to the distinctiveness of their phonological properties, while also being well matched on dimensions relating to semantic associations and visual imagery. In addition, both of the latter indices were below the midpoint of the scales in all cases, suggesting that these items were not strongly linked to semantic associations or vivid mental images.

Design. A repeated measures design was used, with all of the participants receiving the eight target syllogisms, preceded by two single-model problems as practice items. The eight target 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: logic (valid vs. invalid conclusions) and phonological distinctiveness (distinctive vs. nondistinctive contents). The participants were required either to accept or to reject presented conclusions.

Procedure. The participants were presented with the syllogisms in test booklets along with the following instructions:

> This is an experiment to test people's reasoning ability. You will be given 10 problems. On each page, you will be shown two statements describing monsters, and you will be asked whether a conclusion (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, tick 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 (valid vs. invalid) and phonological distinctiveness (distinctive vs. nondistinctive) are presented in Table 2. It is clear that the participants found these syllogisms difficult, as is evidenced by the generally high acceptance rates for conclusions irrespective of logical validity (i.e., the participants accepted many more invalid conclusions than they should have according to logical standards of reasoning). We note, however, that a bias toward acceptance of invalid conclusions is a standard aspect of syllogistic reasoning performance (e.g., Evans et al., 1993) and that the acceptance rates in Experiment 1 are within the normal range associated with multiple-model problems, which are the most difficult of all syllogism types (Johnson-Laird & Byrne, 1991).

Table 2
Percentage of Conclusions Accepted As a Function of Logic and Phonological Distinctiveness in Experiment 1

Logical Status	Phonological Distinctiveness		
	Nondistinctive	Distinctive	Overall
Valid	72	77	75
Invalid	67	61	64
Difference	5	17	11

A Wilcoxon signed-ranks test indicated that significantly more valid conclusions were accepted than invalid ones ($z = 2.44$, $p < .01$). Separate Wilcoxon tests revealed that the effect of logic was reliable for syllogisms with phonologically distinctive contents ($z = 3.24$, $p < .001$) but was not reliable for syllogisms with phonologically nondistinctive contents ($z = 1.24$, $p = .107$).

To confirm the existence of an interaction between logic and phonological distinctiveness, the scores for the invalid problems were subtracted from the scores for the valid problems across participants to give an index of the size of the logic effect for the distinctive versus the nondistinctive contents. A Wilcoxon signed-ranks test demonstrated that the effect of logic differed between these two types of phonological contents in line with the presence of an interaction effect ($z = 2.01$, $p < .05$).

Discussion

In Experiment 1, the use of nonsense terms within syllogisms meant that such terms had no obvious links to known visualizable concepts. As such, the phonological distinctiveness manipulation in the experiment was a relatively pure one, with limited contamination from prior visual or semantic associations. With such controls in place, the results indicated that conclusion evaluation performances were logically superior for the phonologically distinctive problem contents in comparison with the phonologically nondistinctive contents. This evidence appears to support the assumption that phonologically distinctive terms are easier to represent and process during task performance, as was predicted in light of previous demonstrations of phonological similarity effects in working memory (Baddeley et al., 2009).

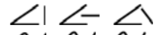
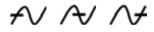
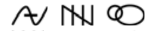
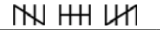
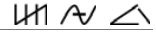
The observation that a phonological distinctiveness manipulation can have an impact on syllogistic performance (improving logical responding for distinctive terms relative to nondistinctive terms) runs counter to the assumption that mental models reflect purely abstract, token-based representations within spatial layouts (e.g., Johnson-Laird, 2005). According to this latter view, surface-level properties of syllogisms, such as the phonology of the presented terms, should have little relevance to the effectiveness of a model-based reasoning strategy. The results of Experiment 1 suggest, therefore, that the representational assumptions of mental models theory may need to be reconsidered. We return to this issue in the General Discussion section after reporting our second experiment.

EXPERIMENT 2

The results of Experiment 1 indicate that distinctiveness effects arising from the phonological properties of presented terms can facilitate syllogistic reasoning within a conclusion evaluation paradigm. In Experiment 2, we aimed to replicate the phonological distinctiveness effect observed in Experiment 1 while also turning our attention to the visual properties of presented syllogistic terms in another experimental condition. Knauff and Johnson-Laird's (2002) visual imagery impedance hypothesis claims that the visual imagery arising from problem contents can hinder people's capacity to construct and use the abstract spatial models that are necessary and sufficient for effective reasoning. Although evidence from studies in which the visualizability of relational information within problems was manipulated supports this hypothesis (e.g., Knauff & Johnson-Laird, 2002; Knauff & May, 2006), it remains unclear whether visual impedance will also arise when the actual *terms* within syllogisms are manipulated. Indeed, as with the phonological distinctiveness effect observed in Experiment 1, it may be that terms that evoke distinctive mental imagery will provide a firmer foundation for syllogistic inference than terms that are visually nondistinctive. Such beneficial effects of visual distinctiveness on reasoning would be in line with the evidence for a visual similarity effect in working memory discussed earlier (e.g., Logie et al., 2000).

To create visually pure terms for use in Experiment 2, we generated bespoke symbols that involved straight lines, wavy lines, angles, and circles (Table 3). These symbols were inserted into syllogisms as terms in the place of written words. By producing such symbol-based syllogisms, we aimed to ensure that prior associations with either phonological or semantic representations were minimized. In addition, because of the use of symbolic materials in one condition alongside phonological materials in another condition, it was possible not only to test Knauff and Johnson-Laird's (2002) visual imagery impedance hypothesis but also to contrast the impact of the visual versus phonological distinctiveness manipulation on reasoning performance. If visual and phonological distinctiveness influence reasoning differently, this would emerge as a three-way interaction between logic (valid vs. invalid conclusions), distinctiveness (distinctive vs. nondistinctive terms), and content (visual vs. phonological). In other words, the expectation would be for the two-way interaction observed with the phonological materials in Experiment 1 to be replicated, whereas a larger, smaller, nonexistent, or reverse two-way interaction would be seen with the visual materials (the latter indicating visual impedance). Conversely, if visual and phonological distinctiveness have equivalent, beneficial influences on reasoning, a two-way interaction between logic and distinctiveness would be present, but there would be no three-way interaction.

Table 3
Symbols Used As Syllogistic Terms in the Visual Condition of Experiment 2

Nondistinctive Visual Content		Distinctive Visual Content	
Symbol Set	Symbols	Symbol Set	Symbols
1		5	
2		6	
3		7	
4		8	

Method

Participants. An opportunity sample comprising 67 female and 42 male participants was tested. The mean age of the participants was 31.1 years (SD = 13.2). None of the participants had taken formal instruction in logic, and all were tested individually.

Materials. The logical forms of the problems in Experiment 2 were identical to those in Experiment 1. Syllogisms contained either phonological or visual terms. The phonological terms were the same one-syllable nonsense words used in Experiment 1. The symbolic terms were simple symbols (see Table 3) comprising two component parts, which we refer to as a *base element* and a *floating element* (e.g., a big oval and a smaller circle, an angle and a small line). The visually nondistinctive syllogisms were those in which the A-, B-, and C-term symbols contained an identical base element, but the relative location or orientation of the single floating element varied among the three terms (see Symbol Sets 1–4 in Table 3). Visually distinctive syllogisms were drawn from the same pool of symbols, but it was ensured that the A, B, and C terms were always distinct from one another (see Symbol Sets 5–8 in Table 3).

To confirm the effectiveness of our visual distinctiveness manipulation, we carried out a pretest using 15 undergraduate students who received payment for their participation. Each participant was given a booklet containing the four sets of distinctive visual terms and the four sets of nondistinctive visual terms, which are depicted in Table 3. Each page of the booklet presented a single set of three symbols with a series of rating tasks below the terms. All of the rating scales were 100-mm horizontal lines with labeled endpoints. The order of the three symbols on each page was independently randomized for each participant, as was the order of the symbol sets within each booklet. The participants were asked to imagine that the symbols denoted the membership of fictitious tribes.

The first rating task was in response to the question "How visually distinctive are these symbols?" with the presented scale ranging from *not at all visually distinctive* to *highly visually distinctive*. As was predicted, the scores revealed a marked separation between the visually distinctive item sets (M = 88.2, SD = 9.8) and the visually nondistinctive item sets (M = 21.3, SD = 17.7) [$F(1,14) = 108.02$, $MS_e = 310.75$, $p < .001$, $\eta_p^2 = .89$].

The second rating task was in response to the question "To what extent do these symbols relate to real symbols that you are familiar with?" with the scale ranging from *not at all related* to *highly related*. The scores here confirmed that the participants viewed neither the visually distinctive item sets (M = 46.0, SD = 16.4) nor the visually nondistinctive item sets (M = 42.6, SD = 19.8) as relating particularly closely to familiar symbols [$F(1,14) = 0.34$, $MS_e = 255.73$, $p = .57$, $\eta_p^2 = .02$].

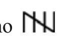
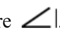
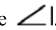
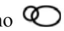
The third rating task requested a response to the question "To what extent do these symbols remind you of words that you are familiar with?" with the scale ranging from *not at all* to *very much*. The scores supported the prediction that neither the visually distinctive item sets ($M = 29.4$, $SD = 15.8$) nor the visually nondistinctive item sets ($M = 29.9$, $SD = 19.7$) were inclined to remind the participants of known words [$F(1,14) = 0.02$, $MS_e = 105.65$, $p = .89$, $\eta_p^2 = .01$].

In the final rating task, we asked the question "To what extent do these symbols allow you to build up vivid mental images of tribal membership categories?" with the scale ranging from *very easy to build up vivid mental images* to *very difficult to build up vivid mental images*. As was anticipated, the visually distinctive item sets afforded significantly better mental imagery ($M = 67.0$, $SD = 19.3$) than did the visually nondistinctive item sets ($M = 48.3$, $SD = 23.7$) [$F(1,14) = 4.76$, $MS_e = 549.10$, $p = .047$, $\eta_p^2 = .25$].

Overall, the pretest data support the view that our novel symbolic terms were effectively polarized in terms of their visual distinctiveness and their capacity to facilitate the construction of vivid mental images of denoted categories. At the same time, the two sets of symbols were well matched on dimensions relating to both known symbols or known words, with measures on these dimensions being uniformly below the midpoint of the respective scales.

Design. A mixed design was used. For one group of participants, the syllogisms had phonological content (see Appendix A for a list of the phonological target problems used), and for the other group, the syllogisms had visual content (see Appendix B for a list of the visual target problems). In addition to this between-participants factor, there were two repeated measures factors: logic (valid vs. invalid conclusions) and distinctiveness (distinctive vs. nondistinctive contents). The participants were required either to accept or to reject the conclusion that was presented with each syllogism. The eight target problems that were given to each participant were presented in a random order. This order was rotated so that each problem appeared once in each serial position, creating eight versions of the test booklet for each type of content. These target problems were preceded by two one-model practice syllogisms.

Procedure. The instructions for the participants who received the phonological syllogisms were the same as those used in Experiment 1. For the visual syllogisms, the following scenario was used to provide the participants with a conceptual basis for the symbolic problem contents:

> This is an experiment to examine people's reasoning ability. Please read the following instructions carefully.
>
> In the Zimporian jungle live many small tribes. Each tribe uses a different symbol to identify its members—for example,
>
>  and ,
>
> Because of marriages between members of different tribes, some individuals are members of more than one tribe. For example,
>
> some  are , and all  are .
>
> However, some tribes do not allow marriages with members of certain other tribes. Consequently,
>
> no  are , and no  are .
>
> You have recently been appointed British Ambassador to Zimporia. It is important, therefore, that you have some practice in using Zimporian tribal symbols and understand the relationships between tribes. To help you with this, you will be given 10 problems. On each page, you will be shown two statements describing the relationships between tribes.
>
> You are asked whether certain conclusions (given below the statements) may be logically deduced from the two statements. You should answer this question on the basis of 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, tick the "false" box. For example,
>
> No  are .
> All  are .
> Therefore, no  are .
> True( ) False( )
>
> 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. Thank you very much for participating.

Results

The percentages of conclusions accepted as a function of content, distinctiveness, and logic are presented in Table 4. A Wilcoxon signed-ranks test showed that, overall, significantly more valid than invalid conclusions were accepted ($z = 2.90$, $p < .01$). Separate Wilcoxon tests revealed that this effect of logic was reliable with the distinctive problem contents ($z = 3.57$, $p < .001$) but was not reliable with nondistinctive problem contents ($z = 1.30$, $p = .19$). To validate the apparent interaction between logic and distinctiveness, the scores for the invalid problems were subtracted from the scores for the valid problems across participants to give an index of the size of the logic effect for the distinctive versus nondistinctive problem contents. A Wilcoxon signed-ranks test indicated that the logic effect differed significantly between the distinctive and nondistinctive contents ($z = 2.11$, $p < .05$).

To test for a three-way interaction among content, logic, and distinctiveness, we computed two-way interaction indices for each problem content by subtracting the logic indices for the nondistinctive problems from the logic indices for the distinctive problems. The two-way interaction indices for the participants receiving the phonological contents did not differ significantly from those for the participants receiving the visual contents ($z = 0.28$, $p = .39$), indicating the absence of a three-way interaction. Note, however, that the logic × distinctiveness interactions for each content type (i.e., phonological or visual) were both reliable ($ps < .05$), confirming that the distinctiveness effect was present in each group separately.

Table 4
Percentage of Conclusions Accepted As a Function of Content, Logic, and Distinctiveness in Experiment 2

	Content					
	Phonological			Symbolic		
Logical Status	Nondistinctive	Distinctive	Overall	Nondistinctive	Distinctive	Overall
Valid	72	78	75	81	82	82
Invalid	68	61	65	74	69	72
Difference	4	17	11	7	13	10

Discussion

The observation of a two-way interaction between logic and distinctiveness in Experiment 2 successfully replicated the results of Experiment 1, which showed a greater logic effect for phonologically distinctive syllogistic contents relative to phonologically nondistinctive contents. Moreover, since the size of the logic × distinctiveness interaction evident with the phonological materials in Experiment 2 did not differ significantly from the size of the same interaction with the visual materials (i.e., there was no three-way interaction), it seems that it is distinctiveness per se that affects syllogistic reasoning performance. In other words, representational distinctiveness has a generic beneficial influence on deductive inference that is not restricted to one particular representational modality.

The results from the visual materials in Experiment 2 also run counter to the visual imagery impedance hypothesis (e.g., Knauff & Johnson-Laird, 2002), since this hypothesis would presumably predict that the mental imagery evoked by distinctive visual contents would hinder people's reasoning with abstract mental models. However, the opposite result was seen to be the case in Experiment 2: Distinctive visual contents led to the emergence of improved logical inference relative to nondistinctive visual contents. This finding concurs with evidence for visually distinctive items' having a positive influence on immediate retrieval from working memory (e.g., Logie et al., 2000).

GENERAL DISCUSSION

In accounting for syllogistic reasoning performance and its inherent biases (e.g., figural effects and conclusion order preferences), the mental models theory assumes that syllogisms—like other deductive problems—are mentally represented as abstract tokens within spatially organized models (e.g., Johnson-Laird, 1996, 1998, 2005). Some of the most compelling evidence supporting the role of abstract spatially based representations in deduction derives from the recent research of Knauff and colleagues using transitive inference problems (e.g., Knauff & Johnson-Laird, 2002; Knauff & May, 2006). This latter work has successfully demonstrated how the mental imagery arising from visually evocative relational terms (e.g., *cleaner than*; *dirtier than*) can slow down people's ability to reason relative to conditions under which relational terms are less visualizable but can nonetheless be envisaged spatially (e.g., *farther north than*; *farther south than*). Knauff and Johnson-Laird suggested that this visual imagery impedance effect arises because a relation such as that which occurs in the premise *the ape is dirtier than the cat* can elicit vivid visual details (e.g., an ape caked with mud) that are irrelevant to the inference. As such, it was proposed that "[i]t will then take additional time to retrieve the information needed to construct the appropriate mental model for making the inference" (Knauff & Johnson-Laird, 2002, p. 370).

Despite this compelling evidence for the abstract spatial basis of mental models in deduction, our research was motivated by the possibility that the visual impedance effect may be limited to cases in which it is the visual properties of *relations* between problem terms that are manipulated, as opposed to the visual properties of the actual terms themselves. Although we agree that the visualizability of relations can engender imagery that is irrelevant to the reasoning task, it nevertheless seemed likely to us that terms that are easier to represent as distinctive, concrete entities could facilitate model construction and reasoning relative to terms that are more difficult to represent in a distinctive visual manner. Likewise, in setting up our research, we also wondered whether terms that have distinctive phonological properties might likewise enable more effective model construction and reasoning than terms that have less distinctive phonological properties, which could make such terms more confusable. The potential for phonological and visual distinctiveness to benefit reasoning has a precedent in research on immediate retrieval from working memory, where it has been established that phonologically or visually distinctive items are more accurately recalled than phonologically or visually similar items (e.g., Baddeley et al., 2009; Logie et al., 2000).

In Experiment 1, we set out to explore whether the phonological distinctiveness of syllogistic terms might have an impact on reasoning effectiveness in a conclusion evaluation paradigm. The findings support the view that phonologically distinctive problem content can enhance reasoning relative to phonologically nondistinctive content. The results of Experiment 2 replicated this phonological distinctiveness effect and also demonstrated an equivalent distinctiveness effect for visually based syllogisms, whereby logical responding was more marked for syllogisms based around distinctive visual terms than for syllogisms that involved nondistinctive visual terms. These latter findings run counter to the visual imagery impedance hypothesis (e.g., Knauff & Johnson-Laird, 2002), instead supporting the view that categorized terms that can be represented as distinct visual entities can facilitate deduction.

Taken together, our results seem to question the idea that categorized terms are necessarily represented within mental models as purely abstract tokens, since such tokens would only be truly abstract if they were amodal and were associated with neither phonological nor visual codes. Our data may instead support the idea that, without distinctive phonological and visual information, individuals will struggle to construct, manipulate, and evaluate the mental tokens that underpin deductive inferences. This is arguably because the representational boundaries between categories remain vague if they are nondistinctive, such that the processing of represented information becomes a muddled endeavor. Indeed, for the syllogisms that contained phonologically or visually nondistinctive terms, the participants appeared to demonstrate difficulty in establishing the validity of presented conclusions, instead showing a bias toward conclusion acceptance, irrespective of logical correctness. In contrast, when distinctive phonological or visual information is available, it appears that this information may clarify the representational boundaries between categories such that reasoning can proceed more effectively.

Our evidence for the involvement of phonological and visual representations in syllogistic inference also concurs with another body of recent research that has examined the role of working memory subsystems in deduction. For example, Gilhooly (2004), in reviewing studies that have manipulated the nature of secondary task loads imposed on reasoners while attempting primary syllogistic tasks, noted that four out of five experiments implicate the involvement of the *phonological loop* subsystem (which is specialized for the representation and processing of phonological information), whereas three out of these five experiments implicate a role for the *visuospatial sketchpad* subsystem (which deals with visually and spatially coded information). Overall, the picture emerging from dual-task studies suggests that multimodal representations may well be associated with syllogistic inference. Again, this view departs somewhat from the assumed primacy of abstract spatially based representations in deduction as espoused by mental model theorists (e.g., Johnson-Laird, 2005).

Interestingly, too, some mental models theorists have recently started to distance themselves from the claim that models entail purely abstract representations. For example, Schaeken, van der Henst, and Schroyens (2006) proposed that reasoners can construct isomeric mental models of presented premises in order to represent indeterminacies and uncertainties. An isomeric model captures all possibilities within a single integrated representation via the addition of concrete, nonspatial elements (i.e., propositional or verbal tags) that can denote uncertainty. A similar notion, espoused by Vandierendonck, Dierckx, and De Vooght (2004), is that of annotated mental models, where annotations are verbal footnotes that act to qualify the meaning of information represented within spatially based models.

Isomeric and annotated models entail rich, multidimensional representations that combine verbal and visuospatial elements within a single integrated format. These recent ideas—when viewed in conjunction with evidence from dual-task studies and our present experiments—lead us to wonder whether the involvement of multimodal information in model-based reasoning may be a typical occurrence in many reasoning contexts, such that reasoners will capitalize on whatever information is available to help with the construction, maintenance, and manipulation of representations during deduction. Sometimes, such multimodal information may lead to reasoning difficulties, as is the case with the impedance arising when visually evocative transitive relations engender imagery that detracts from relational processing. At other times, however, visual and phonological information can facilitate reasoning, as in situations in which the categorized terms referred to in problems are visually or phonologically distinctive.

Notwithstanding the evidence that we have presented, we acknowledge that theorists who are committed to the view that mental models are based on abstract entities could still counter that we have merely demonstrated the benefit of visual and phonological information for *premise processing*, rather than for model-based representation and reasoning, which might still rely exclusively on abstract spatial representations. At first sight, this proposal appears to lead to an unfalsifiable theory, in that whenever evidence is obtained for visual and phonological effects in deduction, these effects can be relegated to an initial premise processing stage, whereas evidence for spatial involvement can be ascribed to a subsequent model-based reasoning stage. Neuroimaging studies may, however, be able to arbitrate successfully on this issue. It could be the case, for example, that early premise processing of visualizable materials activates visual brain areas, whereas subsequent processing that reflects the extraction of abstract mental codes would activate more spatial brain areas. There is, in fact, some evidence supporting this latter position (e.g., Fangmeier, Knauff, Ruff, & Sloutsky, 2006; Knauff, Fangmeier, Ruff, & Johnson-Laird, 2003), although at the moment, it is too early to tell whether such evidence will generalize to a variety of deduction paradigms and content manipulations. We nevertheless agree that neuroimaging research is likely to reveal important insights that will help clarify whether deduction arises through stages of processing that culminate in abstract model-based representations.

This latter (staged) view of reasoning also derives some support from the pioneering studies of spatial reasoning conducted by Mani and Johnson-Laird (1982). In this study, participants were observed to retain resilient verbatim representations of verbally presented *multiple-model* problems (i.e., problems that required two or more mental models for a complete representation of terms and relations), but not for single-model problems that were not open to alternative model-based representations. Mani and Johnson-Laird's evidence suggests that people are highly sensitive to the phonological properties of multiple-model problems, even though inferential processing may itself revolve around subsequently constructed abstract models rather than around initial verbatim traces. We note, too, that all of the syllogistic tasks used in our present research were multiple-model problems, which may, therefore, have demanded some initial maintenance of phonological or visual representations prior to eventual model construction. This initial maintenance of surface level information may provide a locus for the phonological and visual distinctiveness effects that we have observed, while leaving intact the assumption that models themselves are primarily abstract spatially based representations.

Still, it seems valuable to keep sight of alternatives to this staged view of the representations underpinning deduction, especially in light of the recent theorizing discussed previously, which emphasizes the possible role in reasoning of isomeric or annotated models that involve multimodal representations. The possibility that reasoning involves the construction and manipulation of multidimensional representations within a single dynamic storage system that is capable of seamlessly integrating both phonological and visuospatial information seems very attractive to us. At least some of the appeal here derives from the links that we see to interesting developments in the field of working memory research, particularly Baddeley's (2000, 2002) proposals that an episodic buffer may be needed as part of the working memory system in order to provide temporary storage, so as to maintain unitary episodic representations of multidimensional information. Indeed, Baddeley himself draws connections between reasoning and the concept of the episodic buffer when he states that the buffer "allows multiple sources of information to be considered simultaneously, creating a model of the environment that may be manipulated to solve problems and plan future behaviours (Johnson-Laird, 1983)" (Baddeley, 2002, p. 92).

There is clearly much work yet to be done to determine whether syllogistic inference is best explained as involving integrated multidimensional models located within some episodic storage system or as involving abstract amodal spatially based models that are extracted after a stage of initial premise processing. At the very least, our data support the view that distinctive phonological and visual contents can influence the effectiveness of syllogistic inference. As such, we suggest that effects arising from the surface-level features of presented problems necessitate very serious consideration for the derivation of theoretical accounts of the representations that underpin deduction.

AUTHOR NOTE

We thank Natalie Gundry, Richard Nakielny, Karen Matthews, and Rachael Tidswell for their assistance in data collection. We are also grateful for the many valuable comments on previous versions of this article provided by Klaus Oberauer, Kinga Morsanyi, and three anonymous reviewers. Correspondence concerning this article should be addressed to L. J. Ball, Department of Psychology, Lancaster University, Lancaster LA1 4YF, England (e-mail: l.ball@lancaster.ac.uk).

REFERENCES

Bacon, A. M., Handley, S. J., & McDonald, E. L. (2007). Reasoning and dyslexia: A spatial strategy may impede reasoning with visually rich information. *British Journal of Psychology*, **98**, 79-92.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, **18**, 362-365.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, **4**, 417-423.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, **7**, 85-97.

Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Hove, U.K.: Psychology Press.

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., & Johnson-Laird, P. N. (1995). Development of syllogistic reasoning. *American Journal of Psychology*, **108**, 157-193.

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

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

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

Clement, C. A., & Falmagne, R. J. (1986). Logical reasoning, world knowledge, and mental imagery: Interconnections in cognitive processes. *Memory & Cognition*, **14**, 299-307.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, **55**, 75-84.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, **55**, 429-432.

De Soto, L. B., London, M., & Handel, M. S. (1965). Social reasoning and spatial paralogic. *Journal of Personality & Social Psychology*, **2**, 513-521.

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

Espino, O., Santamaría, C., Meseguer, E., & Carreiras, M. (2005). Early and late processes in syllogistic reasoning: Evidence from eye-movements. *Cognition*, **98**, B1-B9.

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

Fangmeier, T., Knauff, M., Ruff, C. C., & Sloutsky, V. (2006). fMRI evidence for a three-stage model of deductive reasoning. *Journal of Cognitive Neuroscience*, **18**, 320-334.

Fleming, P., Ball, L. J., Ormerod, T. C., & Collins, A. F. (2006). Analogue versus propositional representation in congenitally blind individuals. *Psychonomic Bulletin & Review*, **13**, 1049-1055.

Garnham, A., & Oakhill, J. V. (2005). Accounting for belief bias in a mental model framework: Comment on Klauer, Musch, and Naumer (2000). *Psychological Review*, **112**, 509-518.

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

Gilhooly, K. J. (2004). Working memory and reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 49-77). Cambridge: Cambridge University Press.

Hitch, G. J., Halliday, M. S., Schaafstal, A. M., & Schraagen, J. M. C. (1988). Visual working memory in young children. *Memory & Cognition*, **16**, 120-132.

Hue, C., & Ericsson, J. R. (1988). Short-term memory for Chinese characters and radicals. *Memory & Cognition*, **16**, 196-205.

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

Johnson-Laird, P. N. (1996). Images, models and propositional representations. In M. de Vega, M. J. Intons-Peterson, M. Marschark, P. N. Johnson-Laird, & M. Denis (Eds.), *Models of visuospatial cognition* (pp. 90-127). Oxford: Oxford University Press.

Johnson-Laird, P. N. (1998). Imagery, visualization, and thinking. In J. Hochberg (Ed.), *Perception and cognition at century's end* (pp. 441-467). San Diego: Academic Press.

Johnson-Laird, P. N. (2005). Mental models and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 185-208). New York: Cambridge University Press.

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, U.K.: Erlbaum.

Johnson-Laird, P. N., Byrne, R. M. J., & Tabossi, P. (1989). Reasoning by model: The case of multiple quantification. *Psychological Review*, **96**, 658-673.

Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: Behavioral measures and cortical activity. *Journal of Cognitive Neuroscience*, **4**, 559-573.

Knauff, M., & Johnson-Laird, P. N. (2002). Visual imagery can impede reasoning. *Memory & Cognition*, **30**, 363-371.

Knauff, M., & May, E. (2006). Mental imagery, reasoning, and blindness. *Quarterly Journal of Experimental Psychology*, **59**, 161-177.

Knauff, M., & Schlieder, C. (2005). Spatial inference: No difference between mental images and models. *Behavioral & Brain Sciences*, **27**, 589-590.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology*, **53A**, 626-646.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition*, **10**, 181-187.

Miyake, A., & Shah, P. (Eds.) (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* Cambridge: Cambridge University Press.

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.

Schaeken, W., van der Henst, J.-B., & Schroyens, W. (2006). The mental models theory of relational reasoning: Premises' relevance, conclusions' phrasing and cognitive economy. In W. Schaeken, A. Vandierendonck, W. Schroyens, & G. d'Ydewalle (Eds.), *The mental models theory of reasoning: Extensions and refinements* (pp. 129-150). Mahwah, NJ: Erlbaum.

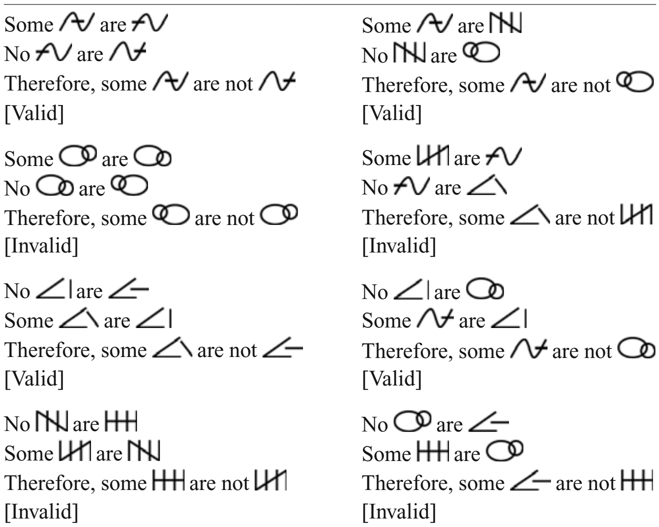
Shaver, P., Pierson, L., & Lang, S. (1975). Converging evidence for

the functional significance of imagery in problem solving. *Cognition*, **3**, 359-375.
STERNBERG, R. J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General*, **109**, 119-159.
STUPPLE, E. J. N., & BALL, L. J. (2005). An inspection-time analysis of figural effects and processing direction in syllogistic reasoning. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (pp. 2092-2097). Alpha, NJ: Sheridan Printing.
STUPPLE, E. J. N., & BALL, L. J. (2007). Figural effects in a syllogistic evaluation paradigm: An inspection-time analysis. *Experimental Psychology*, **54**, 120-127.
STUPPLE, E. J. N., & BALL, L. J. (2008). Belief–logic conflict resolution in syllogistic reasoning: Inspection-time evidence for a parallel-process model. *Thinking & Reasoning*, **14**, 168-181.
VANDIERENDONCK, A., DIERCKX, V., & DE VOOGHT, G. (2004). Mental model construction in linear reasoning: Evidence for the construction of initial annotated models. *Quarterly Journal of Experimental Psychology*, **57A**, 1369-1391.

APPENDIX A
Target Syllogisms Used in Experiment 1, Showing the Logical Status of Presented Conclusions

Nondistinctive Phonological Content	Distinctive Phonological Content
Some bubs are bebs No bebs are babs Therefore, some bubs are not babs [Valid]	Some zaps are toks No toks are yugs Therefore, some zaps are not yugs [Valid]
Some fuds are fods No fods are feds Therefore, some feds are not fuds [Invalid]	Some fubs are haps No haps are beks Therefore, some beks are not fubs [Invalid]
No herks are horks Some harks are herks Therefore, some harks are not horks [Valid]	No harks are paps Some fids are harks Therefore, some fids are not paps [Valid]
No juks are jeks Some jiks are juks Therefore, some jeks are not jiks [Invalid]	No keps are yogs Some zucks are keps Therefore, some yogs are not zucks [Invalid]

APPENDIX B
Target Syllogisms Used in Experiment 2, Showing the Logical Status of Presented Conclusions



Nondistinctive Visual Content	Distinctive Visual Content
Some are No are Therefore, some are not [Valid]	Some are No are Therefore, some are not [Valid]
Some are No are Therefore, some are not [Invalid]	Some are No are Therefore, some are not [Invalid]
No are Some are Therefore, some are not [Valid]	No are Some are Therefore, some are not [Valid]
No are Some are Therefore, some are not [Invalid]	No are Some are Therefore, some are not [Invalid]

(Manuscript received March 21, 2008;
revision accepted for publication February 24, 2009.)