Memory & Cognition
2006, 34 (1), 196-206

Effects of instruction presentation mode in comparative judgments

SAMUEL SHAKI, CRAIG LETH-STEENSEN, and WILLIAM M. PETRUSIC
Carleton University, Ottawa, Ontario, Canada

In each of two experiments, the comparative instructions in a symbolic comparison task were either varied randomly from trial to trial (mixed blocks) or left constant (pure blocks) within blocks of trials. In the first experiment, every stimulus was compared with every other stimulus. The symbolic distance effect (DE) was enhanced, and the semantic congruity effect (SCE) was significantly larger, when the instructions were randomized than when they were blocked. In a second experiment, each stimulus was paired with only one other stimulus. The SCE was again larger when instructions were randomized than when they were blocked. The enhanced SCE and DE with randomized instructions follow naturally from evidence accrual views of comparative judgments.

Any complete theory of the process of comparing either perceptual or remembered stimuli must provide an explanation for the *semantic congruity effect* (SCE). The SCE is characterized by an interaction between the particular comparative instruction required and the location of the stimulus pair to be discriminated on the underlying continuum. For example, as in the landmark psychophysical experiments of Audley and Wallis (1964) that brought the SCE to the attention of contemporary psychophysicists, the time to select the darker of two relatively dark lights is shorter than the time to select the brighter. Conversely, selection time of the brighter of two relatively bright lights is shorter than the selection time of the darker.

In Audley and Wallis's (1964) experiment, and in the replication and extension of Wallis and Audley with the pitch dimension, the comparative instructions (i.e., select the darker or select the lighter) were presented separately in counterbalanced blocks. In the large number of ensuing experiments, both in the strictly perceptual domains (e.g., Marschark & Paivio, 1981; Petrusic, 1992; Petrusic & Baranski, 1989) and in the voluminous literature with symbolic comparisons (for reviews, see, e.g., Banks, 1977; Leth-Steensen & Marley, 2000), both *blocked* and *randomized* modes of presentation of the instructions have been used. Curiously, it remains unclear whether instruction presentation mode affects the magnitude of the SCE.

Shaki and Algom (2002) have previously argued that contrasting conditions in which the instructions are blocked, as compared with when they are randomly intermixed over trials, should provide a strong test of Leth-Steensen and Marley's (2000) connectionist, *instructional pathway interference model*, because (for reasons to be detailed later) this model would almost certainly predict enhanced SCEs for randomized instructions, as compared with blocked ones. More generally, though, it is of interest to determine the theoretical implications that contrasting the SCE over instruction presentation modes have for most of the currently popular alternative theories of the SCE.

Another important and robust comparative judgment phenomenon is the symbolic distance effect (DE). This effect is characterized by longer response times (RTs) for comparisons of stimuli that are closer in magnitude than for comparisons of stimuli with larger differences in magnitude. Because DEs are invariably assumed to arise out of the same decision processes that give rise to the SCE (see Leth-Steensen & Marley, 2000, for a full discussion of theoretical accounts for the DE), it is also of interest to determine whether or not randomized and blocked instruction presentation modes have parallel effects on the sizes of both the symbolic DE and the SCE. Hence, in the present experiments, we directly examined these issues by comparing blocked versus randomized instruction presentation conditions in comparative judgments, using the classic Moyer (1973) stimuli of names of animals varying in size.

EXPERIMENT 1

Method

Participants

Sixteen Carleton University students participated in one 80-min session to satisfy course requirements. All the participants reported normal or corrected-to-normal vision.

Stimuli and Design

Six animal names, all three-letter words in English, printed in Times New Roman font (25 point, bold) defined the stimulus set. Three names were of relatively small animals (ant, bee, and rat), whereas the other three names were of relatively large animals (cat,

The work was supported by Natural Sciences and Engineering Research Council of Canada grants to W.M.P., C.L.-S., and Lise Paquet. Address correspondence to S. Shaki, Department of Behavioral Sciences, College of Judea and Samaria, Ariel 44837, Israel (e-mail: samuel_shaki@hotmail.com), to C. Leth-Steensen, Department of Psychology, Carleton University, Ottawa, ON, K1S 5B6 Canada (e-mail: craig_leth_steensen@carleton.ca), or to W. M. Petrusic, Department of Psychology, Carleton University, Ottawa, ON, K1S 5B6 Canada (e-mail: bill_petrusic@carleton.ca).

hog, and cow). Each of the six animal names was paired with every other animal name. Each of these 15 pairs was presented in each of the two possible left–right position orders, resulting in 30 stimulus pairs in the design.

The two forms of the comparative instructions (“smaller” and “larger”), printed in David font (30 point, bold), occurred equally often with each stimulus pair and were varied randomly from trial to trial on half of the blocks but were constant over a block for the other half. Each block consisted of three replications of the 30 stimulus pairs (in subblocks of 30 trials). Each block of the preceding practice trials consisted of one replication of the 30 stimulus pairs.

Both the order in which the blocked and the randomized conditions were presented and the order in which the two instructions were used in the blocked conditions were counterbalanced through the assignment of 4 participants to each of four groups. The four groups were defined by the following sequences of blocks of trials, with R denoting a block of randomized instruction trials, S a block with the instruction to choose the smaller, and L a block with the instruction to choose the larger: SLRRRRLS, LSRRRRSL, RRSLLSRR, and RRLSSLRR, respectively. Precisely the same sequence of blocks of trials was used for the practice trials as for the experimental trials for each group. The participants were not aware of the partition into practice and experimental trials. The order of presentation of the stimulus pairs (and instructions in the randomized condition) within blocks was random and was different for each participant.

Procedure

The participants were tested individually in a dimly lit room, seated approximately 60 cm from the center of the video monitor. The participants were told that the presentation of the comparative instruction word served as a warning for the next trial and indicated whether they were to choose either the smaller or the larger animal in the pair. After an additional 750 msec, the pair of animal names appeared while the comparative instruction remained on the screen. The participants’ task was to press the mouse button[1] on the same side as that on which the smaller (or the larger, respectively) member of the pair of animal names appeared. The presentation of the stimuli and the comparative instruction were response terminated. The next trial began 1,000 msec later. The participants were encouraged to respond quickly but accurately. The 80-min session included five planned breaks, each of which ended with the participants’ decision to continue.

The pairs of animal names appeared at the respective centers of the left and right hemifields on the white background of a 17-in. (43-cm) ViewSonic video monitor, and the comparative instructions appeared at the center of the upper third of the screen. Event sequencing, randomization of trials and instructions, and recording of responses and RTs were under the control of SuperLab software run on a Pentium III microprocessor.

Results

The findings are presented in two main sections. The first section presents RT analyses, and the second section focuses on error rates. For each participant, in all the analyses, the dependent variables are the mean RTs for correct responses and mean arcsine-transformed proportions of errors in each cell of the design. Appendices A and B present mean RTs, and standard deviations for correct and for error responses for each cell of the design for the blocked and the randomized conditions, respectively. In each ANOVA, the Huynh–Feldt epsilon adjustment was used, although the degrees of freedom reported are those defined by the design.

RT Analyses

DEs. The present design permits full examination of DEs, given that each of the six stimuli was compared with every other stimulus. Figure 1 plots mean RTs as a function of the ordinal distance separating the two stimuli in a pair for the randomized and for the blocked conditions. An ANOVA with the two instruction presentation modes (blocked and randomized) and distance (ordinal steps of 1, 2, 3, 4, and 5) as within-participants factors was conducted. As was expected, robust DEs were obtained overall [$F(4,60) = 90.20, p < .001$]. Figure 1 also shows that the pairs in the blocked condition were compared more quickly, overall, than the pairs in the randomized condition. On average, RTs were 859 and 1,026 msec in the blocked and the randomized conditions, respectively, and this main effect of instruction presentation mode was statistically reliable [$F(1,15) = 17.03, p < .001$]. Importantly, the linear component of the interaction of instruction presentation mode and distance was significant [$F(1,15) = 12.72, p < .003$]. As is evident in Figure 1, the DE was somewhat enhanced with the randomized instructions.

SCEs. In order to examine the SCE, a subset of the data, pairs separated by a single ordinal unit, was reanalyzed. Thus, another ANOVA was conducted on these data with instruction presentation mode (blocked and randomized), stimulus pair (five adjacent pairs), and instruction type (“smaller” and “larger”) as within-participants factors.

As panel A in Figure 2 shows, SCEs are clearly evident with both the randomized and the blocked instructions

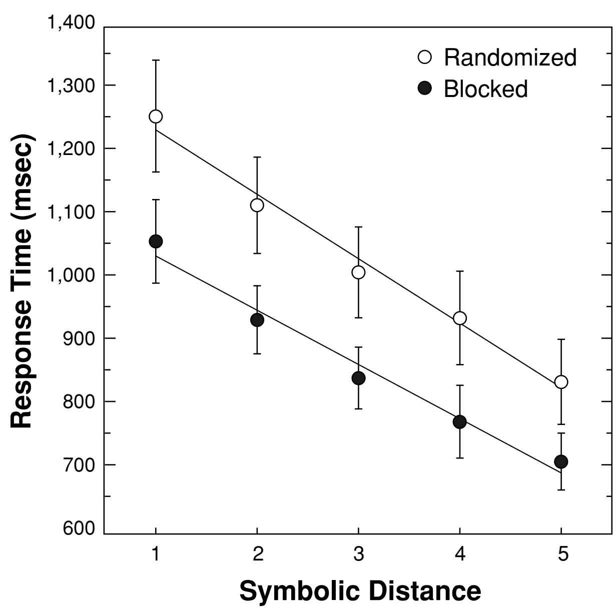

Figure 1. Mean response times as a function of symbolic distance (in ordinal units) for the blocked and the randomized instruction conditions in Experiment 1. Plots of the least squares linear regressions are also provided for each condition.

for these adjacent pairs, and the overall interaction of stimulus pair and instruction type was statistically reliable [$F(4,60) = 29.83$, $p < .001$]. More important, significantly larger SCEs were obtained with the randomized instructions than with the blocked instructions, evident in the statistically reliable three-way interaction involving instruction presentation mode, stimulus pair, and instruction type [$F(4,60) = 3.14$, $p < .023$]. For the smallest stimulus pair, the size of the SCE was 407 msec with the randomized instructions (i.e., the participants selected the smaller animal in the pair 407 msec more quickly than they selected the larger animal) but only 314 msec with the blocked instructions. Similarly, for the largest pair, the size of the SCE was 485 msec with the randomized instructions (i.e., participants selected the larger animal in the pair 485 msec more quickly than they selected the smaller animal in the pair) but only 308 msec with the blocked instructions.

Panel B in Figure 2 provides an alternative, more direct view of the effects of instruction presentation format on the magnitude of the SCE. These plots provide an index of the SCE, based on subtraction of the RTs with the "smaller" instruction from the RTs with the "larger" instruction. As is evident, linear regressions (the adjacent pairs were coded in steps of size 1) provide a convenient summary of the SCE, and, importantly, the slope of the SCE index in the randomized instruction condition (−220 msec/step) is greater than that in the blocked condition (−150 msec/step). Taken together, these plots of the SCE index provide an alternative view of the basis for the significant three-way interaction involving instruction presentation mode, stimulus pair, and instruction type.

Overall, the speed advantage for the blocked condition over the randomized condition is evident in this subset of data, as it is in the full set of data in Figure 1. Mean RTs were 1,256 msec with the randomized instructions and 1,061 msec with the blocked instructions [$F(1,15) = 15.34$, $p < .001$]. Finally, the classic inverted-U-shaped end effect was also obtained. Performance was faster, overall, for the smallest (1,108 msec) and largest (999 msec) stimulus pairs than for the three pairs of intermediate size (1,257 msec), and the overall main effect of stimulus pair was statistically reliable [$F(4,60) = 9.01$, $p < .001$].

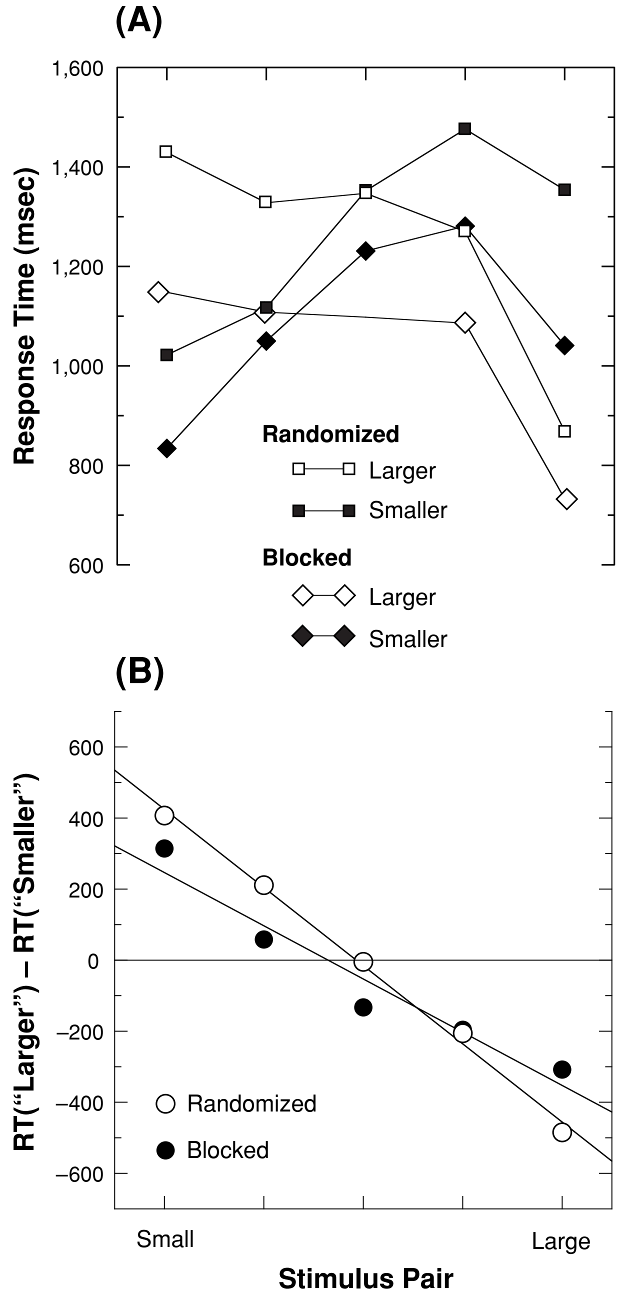

Figure 2. (A) Mean response times (RTs) for the adjacent stimulus pairs with each instruction type in the randomized and the blocked instruction conditions in Experiment 1. (B) The semantic congruity effect index, defined by RT("Larger") − RT("Smaller"), for each of the adjacent stimulus pairs in the randomized and the blocked instruction conditions. Plots of the least squares linear regressions, assuming equal spacing of adjacent stimulus pairs, are also provided for each condition in panel B.

Error Analyses

DEs. The participants made 1.85% errors overall in the blocked instruction condition and 3.04% errors in the randomized instruction condition [$F(1,15) = 17.21$, $p < .001$]. The main effect of distance was highly reliable [$F(4,60) = 18.31$, $p < .001$], with, as was expected, a higher error rate for the one-step pairs (6.4%) than for the five-step pairs (1.1%).

SCEs. For the adjacent pairs, the participants made 6.6% errors with the blocked instructions and 6.3% errors with the randomized instructions ($F < 1$). Paralleling the end effect in the RTs, the participants made fewer errors with the smallest (7.2%) and with the largest (2.9%) pairs than with the intermediate size (8.5%) pairs. This main effect of stimulus pair was statistically reliable [$F(4,60) = 6.22$, $p < .001$]. No other main effects or interactions attained statistical significance. Thus, the main thrust of these error analyses is to show that the previous RT findings are *not* a consequence of differential speed–accuracy trade-offs.

Discussion

The empirical findings of Experiment 1 are clear. The SCE is significantly larger when the instructions randomly vary from trial to trial than when they are constant over a block of trials. As well, the symbolic distance effect is enhanced when instructions are randomized.

However, there is an aspect of the design in Experiment 1 that permits a somewhat artifactual explanation of the three-way interaction involving the SCE to be entertained. Note that to permit full examination of the symbolic DE, each animal name was paired with every other animal name. As a consequence, each animal name, except the smallest and the largest names, was presented with a smaller member in some of the pairs and with a larger member in the other pairs. Hence, for the randomized instructions, many of the stimuli would have been responded to and, hence, interpreted as being the larger stimulus on some trials and as being the smaller stimulus on other trials. According to Logan's (1988, 1990) instance theory, whenever a previously encountered item is presented on a new trial, the current interpretation of that stimulus will be influenced by an automatic retrieval of the previous interpretation involving that stimulus. Overall, the retrieval of such instance information should lead to interference (i.e., negative priming) when one is switching between randomized instructions that does not arise when the same instructions are used throughout a block (see also MacDonald & Joordens, 2000; Strayer & Grison, 1999; Wood & Milliken, 1998). This point is particularly relevant to comparisons involving semantically incongruent end pairs of stimuli, such as choosing the larger of the pair (ant, bee) in the randomized instruction presentation condition in Experiment 1, because in every other pair in which the correct stimulus item (e.g., bee) for the semantically incongruent comparison is a member, that same stimulus item would always have been identified as being the correct choice for the opposite instruction.

In Experiment 2, we attempted to avoid this possibility by pairing each stimulus item with only one other item. Thus, all instances involving a member of a pair should lead to the same interpretation, and Logan's (1988, 1990) instance theory would not apply in this context. If the findings in Experiment 1 were simply a consequence of pairing each stimulus with every other stimulus (vis-à-vis Logan's instance theory), the mode of instructional presentation should have no effect on SCE magnitude when each stimulus is paired with a single unique stimulus.

EXPERIMENT 2

Method

Participants

Sixteen Carleton University students participated in a 1-h session to satisfy course requirements. All the participants reported normal or corrected-to-normal vision.

Stimuli and Design

The size norms of Paivio (1975) were used to select 16 animal names: eight relatively small animals (flea, snail, bee, crab, frog, mouse, rat, and dove) and eight relatively large animals (dog, goat, wolf, lion, cow, horse, bear, and whale). Four relatively small animal pairs (bee–rat, flea–crab, frog–dove, and snail–mouse) and four relatively large animal pairs (dog–cow, goat–lion, wolf–bear, and horse–whale) were created. As is evident, each stimulus appeared but once in the stimulus pair set and was unique to each stimulus pair. As well, both members of each stimulus pair included the same number of letters.

Each participant received eight experimental blocks, with 64 trials in each block, preceded by eight blocks of 16 practice trials each. As in Experiment 1, the two forms of the comparative instructions occurred equally often and were varied randomly from trial to trial on half of the blocks but were constant over a block for the other half. In the randomized condition, the 64 trials arose from replicating twice the factorial combination of eight stimulus pairs by two left–right position orders by two instructions. In the blocked condition, the 64 pairs arose from replicating four times the factorial combination of eight stimulus pairs by two left–right position orders. Both the order in which the blocked and the randomized conditions were presented and the order in which the two instructions were used in the blocked conditions were counterbalanced through the assignment of 4 participants to each of four groups in the same manner as that described in Experiment 1. The order of presentation of the stimulus pairs (and instructions in the randomized condition) within blocks was random and different for each participant.

Procedure

The procedure was essentially the same as that in Experiment 1, except for the fact that the participants were now provided with a short, self-terminated break after every 128 trials.

Results

As in Experiment 1, the findings are presented in two main sections involving separate analyses of the mean correct RTs and arcsine-transformed mean proportions of errors in each cell of the design (with the Huynh–Feldt epsilon adjustment of the degrees of freedom in the ANOVAs).

RT Analyses

Figure 3 provides plots of mean RTs for the four small and for the four large pairs, with each instruction type, separately for the blocked and for the randomized conditions. As is shown, the relatively small pairs were compared more quickly with the instruction to choose the smaller stimulus than with the instruction to choose the larger, whereas this finding was reversed for the relatively large pairs. An ANOVA with the two stimulus pair categories (small and large, obtained after combining RTs over the four stimulus pairs in each category), the two instruction types, and the two instruction presentation modes as within-participants factors showed the overall SCE to be statistically reliable [$F(1,15) = 31.41$, $p < .0001$]. Importantly, the three-way interaction involving instruction presentation mode, pair size, and instruction type was also reliable [$F(1,15) = 16.17$, $p < .001$], providing a clear replication of the findings of Experiment 1. As the plots in Figure 3 show, the SCE was larger with the randomized instructions than with the blocked instructions.

Moreover, the increase in SCE with the randomized instructions was not unique to a particular stimulus pair. Rather, as Figure 4 shows, the reduction in the SCE in the blocked condition was evident with all four of the large pairs and with three of the four small pairs.

Overall, the speed advantage for the blocked condition over the randomized condition found in Experiment 1 was

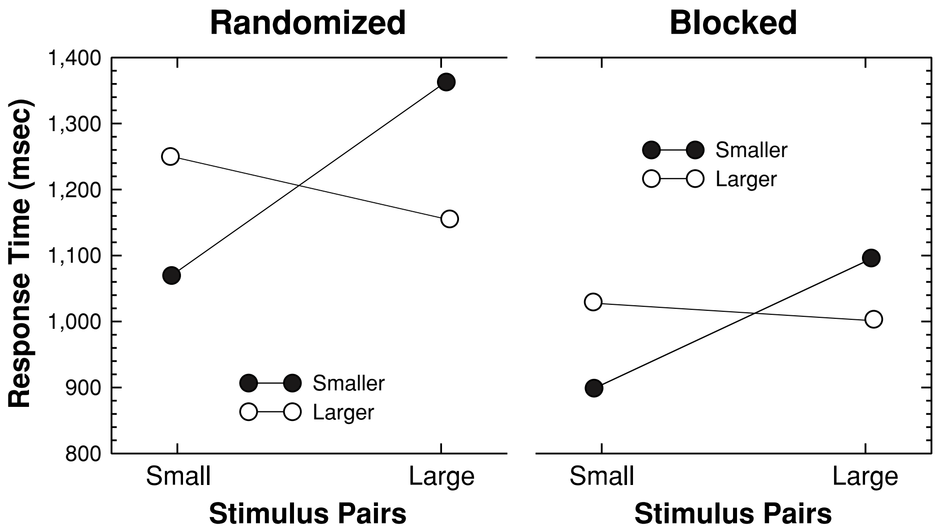

Figure 3. Mean response times with the set of small and the set of large stimulus pairs with each instruction in the randomized and the blocked instruction conditions in Experiment 2.

also evident in this experiment. Mean overall RTs were 1,210 msec with the randomized instructions and 1,006 msec with the blocked instructions [$F(1,15) = 10.33$, $p < .01$]. None of the other main effects or interactions was significant. Unfortunately, given the design, results involving the symbolic DE were not available from this experiment.

Error Analyses

The participants made fewer errors in the blocked condition (2.7%) than in the randomized condition (3.9%), mirroring the pattern obtained for RTs [$F(1,15) = 11.06$, $p < .005$], thereby indicating an absence of a speed–accuracy trade-off effect. No other main effects or interactions attained statistical significance.

Discussion

The results in Experiment 2 are also clear in showing that the mode of presentation of the instructions influences the magnitude of the SCE (i.e., that it is significantly larger when the instructions vary randomly from trial to trial than when they are constant over a block) and provide an important replication and extension of the findings obtained in Experiment 1. Moreover, they indicate that this finding cannot simply be explained away as being due to negative priming.

GENERAL DISCUSSION

The finding of enhanced SCEs in the randomized instruction presentation mode clearly resolves the empirical issue raised by Shaki and Algom (2002) (which was the initial impetus for the present research) and is consistent with predictions based on the model of Leth-Steensen and Marley (2000). Leth-Steensen and Marley's model is a connectionist-based, evidence accrual model that posits the continuous accumulation of information about both the difference in stimulus magnitude and the end-point status of each stimulus item. This information is assumed to be accumulated simultaneously within two competing instructional pathways that are associated with both the relevant and the irrelevant comparative instructions, respectively. Moreover, Leth-Steensen and Marley assume that the irrelevant pathway is *selectively* attenuated, in the spirit of Cohen, Dunbar, and McClelland's (1990) connectionist model of the Stroop phenomenon. In that model, either word-naming or color-naming task-relevant pathways are assumed to be selectively activated according to the situational task demands.

The main aspect of Leth-Steensen and Marley's (2000) model that results in the SCE is the assumption that the strengths of each instructional pathway and, hence, the overall level of competition between them are assumed to be dynamically modulated by the relative location of the stimulus items. For example, a pair of relatively small stimuli would serve to enhance the strength of the instructional pathway associated with the instruction to choose the smaller stimulus and would weaken the strength of the instructional pathway associated with the instruction to choose the larger stimulus. In this example, such dynamic modulation of the competing pathway strengths would then facilitate the process of choosing the smaller item and hinder the process of choosing the larger item (i.e., lead to an SCE). As Leth-Steensen and Marley indicated, this notion of competition between instructional pathways is entirely consistent with the notion of semantic interference originally proposed by Banks and Root (1979). Furthermore, because both instructional sets must be maintained (and used) in the randomized instruction presentation mode, as opposed to only one of the instruction sets in the blocked mode, it follows that selective attenuation of the irrelevant instructional pathway would be less precise for randomized instructions than for blocked instructions. Hence, this model would predict increased

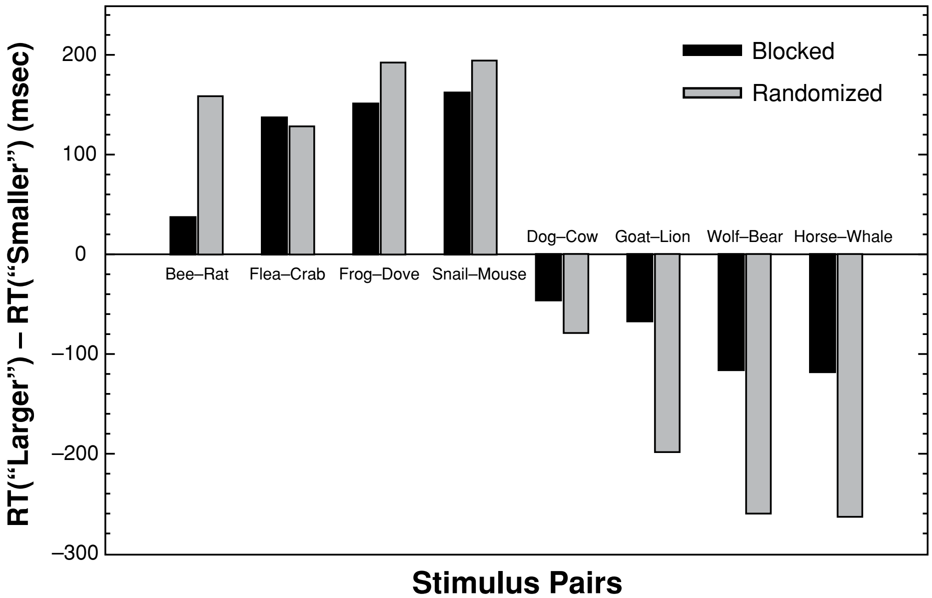

Figure 4. Semantic congruity index, defined by RT("Larger") − RT("Smaller"), for each of the eight stimulus pairs for the blocked and the randomized instruction conditions in Experiment 2.

instructional pathway competition and, consequently, increased SCE when instructions are randomized.

Moreover, some quantitative simulation work with Leth-Steensen and Marley's (2000) model indicates that an explicit prediction of the model is that there should be parallel effects of randomized and blocked instruction presentation modes on the SCE and the symbolic DE. In this work, two versions of the model were run. In one version (i.e., randomized instructions), the parameter specifying the degree of competition between the two instructional pathways was set higher than in the other version (i.e., blocked instructions), but the remaining set of model parameters was kept constant. A full set of model parameters were located that provided simulated RT results that were very similar to those shown for the endpoint pairs in Figure 2 (i.e., with more of an SCE for the randomized than for the blocked instructions). As well, the DEs obtained from these simulation results were somewhat larger for the randomized than for the blocked instructions.

One important additional aspect of these model simulations was that the overall increase in RT for the randomized instruction conditions, in comparison with the blocked instruction conditions, was about half the size of the actual increase observed in the empirical data (indicating that any modeling of the actual data with this model would likely also need to include an additional 80–100 msec or so constant increase in RT for the randomized instructions). Such overall slowing for the randomized instruction condition is directly analogous to the well-known *mixing costs* in RT (Los, 1996) that occur whenever a condition in which some particular stimulus-based or task-based factor is manipulated within blocks of trials (i.e., mixed blocks) is contrasted with a condition in which that same factor is manipulated across blocks of trials (i.e., pure blocks).[2]

Other Theories for the SCE

Evidence-accrual-based theories. Petrusic (1992) has presented strong empirical support for the notion that the SCE occurs at the level of each accrual event within a discrete evidence accrual process (i.e., a slow- and fast-guessing discrete accumulator). That is, the duration of each evidence accrual event is longer for comparisons involving semantically incongruent stimuli than for those involving semantically congruent stimuli. Petrusic further hypothesized that the slowing of the accrual process for semantically incongruent stimuli arises because information regarding the relative magnitudes of the stimuli is of much poorer quality when their locations are incongruent with the form of the comparative instruction.

Alternatively, according to differential *bias* theories of the SCE (Birnbaum & Jou, 1990; Link, 1990, 1992; Schwarz & Stein, 1998), this effect rises as a consequence of a dynamic, strategic adjustment of decisional criteria within an evidence accrual decision process (i.e., a random walk). For example, if the instruction is to choose the smaller stimulus and the stimulus pair contains a very small stimulus, bias theories assume that individuals recognize that this stimulus is likely to be the correct choice and lower the decision criterion associated with the accumulation of evidence for that stimulus (or conversely, recognize that this stimulus is unlikely to be the correct choice when the instruction is to choose the larger and raise that same decision criterion; see also Link, 1992, pp. 172–178).

For both of these theories, any manipulation that slows the overall evidence accrual process should also exaggerate the size of the SCE (in addition to both the overall RTs and the size of the DE, because within these theories, this effect is assumed to arise because more accruals are required to compare stimuli that are closer to one another than to compare those that are farther apart). Hence, one implication of the finding of enhanced SCEs (and DEs) in the randomized instruction condition is that it follows naturally from the notion that randomly mixing the instructions generally serves to slow the evidence accrual process (note that, as far as we are aware, this notion is a novel one that has not explicitly been raised before with respect to task-based mixing costs). One way in which such slowing could be assumed to occur is that when the comparative instructions vary randomly from trial to trial, memory access to the relevant instruction is slower than when the instructions are blocked. If it is also the case that the relevant instruction must be accessed throughout the evidence accrual process, this process will be slower when instructions are randomized than when they are blocked. (In fact, this account is analogous to the one just described for Leth-Steensen and Marley's [2000] model, although the actual mechanisms through which both such overall slowing and enhancement of the SCE would occur in the randomized instruction condition are more precisely specified in that model.)

However, one alternative theoretical account of RT mixing costs is that the greater trial-by-trial uncertainty in mixed blocks regarding the identity of either the stimuli or the task (i.e., the nature of the comparative instructions, in the present case) induces a strategic increase in the overall level(s) of the decision criteria, in order to accommodate an anticipated increase in the processing demands associated with this uncertainty (Los, 1996). Because increasing the decision criteria increases the amount of evidence required by the accrual process, it would also serve to exaggerate any RT effects that are present, such as the SCE (and the DE), within the present paradigm. One additional consequence arising from this criterion adjustment notion, though, which is not supported by the present set of empirical findings, is that raising decision criteria when the comparative instructions are randomized should also likely result in corresponding *decreases* in errors (because such an adjustment necessarily invokes a form of speed–accuracy trade-off).

Non–evidence-accrual-based theories. There are a number of other available theoretical accounts for the SCE that are not specifically evidence accrual based. For example, Banks's (1977; see also Banks, Clark, & Lucy, 1975; Banks & Flora, 1977; Banks, Fujii, & Kayra-Stuart, 1976; and more recently, Cech, 1995; Cech & Shoben, 1985; Cech, Shoben, & Love, 1990) *semantic-coding* theory provides a full and compelling account of comparative judgments, especially with symbolic stimuli. According to semantic-coding theory, whenever a stimulus pair is presented for comparison, each element in the pair is coded categorically. For example, if a relatively small stimulus pair is presented, the elements in the pair might be coded as small and very small—represented more formally as S and S+, respectively. Similarly, a relatively large pair might be coded as L+ and L++ (i.e., very large and extremely large, respectively). Presentation of a particular instruction initiates a memory search for the stimulus with more instances of the code specified by that instruction. For example, given the former case, presentation of the instruction to choose the smaller stimulus in the pair leads to a search for the stimulus element with more of the S code, and hence, the stimulus coded as S+ can be chosen directly. On the other hand, for this same case, presentation of the instruction to choose the larger stimulus in the pair leads to a search for the stimulus with more of the L code, and because the stimuli are coded as S and S+, this search initially fails. Hence, time must be taken to recode the stimuli as L+ and L, and it is this recoding process that is assumed to give rise to the SCE. However, precisely the same stimulus and instructional codes should become activated when the instructions are constant over a block of trials as when they vary randomly from trial to trial. Consequently, precisely the same code search and translation processes should occur for both instructional presentation modes, and the magnitude of the SCE would not be expected to differ.

Alternatively, Marschark and Paivio (1979, 1981) and Kosslyn, Murphy, Bemesderfer, and Feinstein (1977) have developed variants of the *expectancy* theory view of the SCE, primarily in the context of symbolic comparisons. The essence of the expectancy idea is that the instruction directs (much as in semantic priming; see, e.g., Meyer & Schvaneveldt, 1971; Neely, 1977) the memory search for the relevant magnitude features of the to-be-discriminated stimulus pair toward the end of the attribute continuum specified by the instruction. Hence, whenever the stimulus pair location is congruent with the instruction, the search process is semantically facilitated, and when it is not, the search process must be redirected toward the opposite end of the attribute continuum, slowing the comparison process. However, precisely the same expectancy priming processes should occur under conditions in which the instructions are blocked as under conditions in which they are randomized. Consequently, according to expectancy theory, the SCE would also not be expected to vary with instruction presentation mode.

In addition, *reference point* theories (Dehaene, 1989; Holyoak, 1978; Jamieson & Petrusic, 1975; Marks, 1972) posit that stimuli are represented on an analogue continuum and that presentation of an instruction activates an extreme point on the continuum, referred to as a *reference point*. Comparison is based on the ratio of the distances of the representations of the stimuli from the activated reference point, and RTs are assumed to vary inversely with the difference between the ratio of distances and a criterion value (typically, 1 in the unbiased case). This difference between the ratio of distances and the criterion increases with nearness of the stimulus pair to the reference point. In essence, reference point theories assert that stimulus pair discriminability is better the closer the pair is to the activated reference point. However, because precisely the same reference points should become activated when instructions are randomized as when they are blocked, reference point theory also would predict that the SCE effect should be the same for the two instruction presentation modes.

With respect to these three theories, a *constant* RT benefit with blocked instructions could occur either in the semantic-coding and expectancy models or in the reference point model, if it is assumed that upon the appearance of the stimulus pair, individuals are more fully prepared to invoke the relevant memory search or use the relevant reference point, respectively, when the instructions are blocked than when they are randomized. Such constant RT benefits could also occur if the use of blocked instructions somehow speeds some aspect of the processing taking place at either the initial encoding or the final response stage (Los, 1996).

In addition, as has been discussed by Los (1996), responding in blocked conditions could benefit from repeatedly engaging the same cognitive processes (i.e., a process repetition effect, which Los suggested might arise due to the residual activation of pathways that have recently been utilized). In the present case, the cognitive processes being repeated (and hence, speeded) in the blocked instruction mode would be those involved in responding according to only one of the possible comparative instructions. One ramification of this notion is that it is possible to cast both the instructional interference (or actually, the lack thereof in the blocked instruction mode) assumptions of Leth-Steensen and Marley (2000) and the differential accrual rate assumptions of the two other evidence accrual models in terms of such process repetition benefits. The other ramification is that for the reference point and, likely, the expectancy view as well, performance benefits due to process repetition effects would still not be expected to lead to differential SCEs in the two instruction presentation modes. However, with respect to the semantic-coding view, it is possible to envision a scenario in which repeatedly engaging (and hence, speeding) the processes associated with only one of the instructions could lead to a reduction in the SCE, due to the fact that only one version of the code translation process needs to be invoked whenever the instructions are blocked (e.g., L-to-S code translations for the instruction to choose the smaller item).

Finally, an additional point that is relevant here is that it is generally accepted in the dual-task and task-switching literature (e.g., Pashler, 1994; Rogers & Monsell, 1995) that there are performance costs associated with preparing and maintaining two simultaneous task sets, in comparison with maintaining a set to perform only a single type of task (which is directly analogous to the differential instructional set requirements of the randomized and blocked instruction presentation modes, respectively). One way to conceptualize the locus of this cost is that the task sets must be maintained in working memory, and the cost of doing so in terms of limited cognitive resources (i.e., the mental load) is higher for two task sets than for one. Hence, if more resources are devoted to maintaining the task sets, it could be assumed that fewer resources are available to the actual performance of the task, which could slow overall RTs and also potentially enhance the size of any additional RT effects that are present (note that as this assumption is cast here, it does not necessarily require the involvement of any evidence accrual mechanisms). One problem, though, with such an assumption is that there is no a priori way to determine whether the resource requirements of maintaining the instructions and performing the actual comparison process do indeed overlap. Furthermore, there has lately been something of a backlash against the utility of cognitive resources as a theoretical construct, given that almost any kind of empirical effect can be explained away in terms of some kind of differential resource allocation (Sanders, 1997; and note that in a recent extensive theoretical discussion of mixing costs by Los [1996], the construct of limited shared resource capacity was mentioned only in passing).

CONCLUSION

The present results are entirely consistent with theories of the SCE that conceptualize the decision processing in symbolic comparison in terms of the accumulation of evidence. As has been discussed, such theories can provide a natural explanation for the presence of an enhanced SCE in the randomized instruction presentation condition by assuming that the overall accumulation of evidence is slowed when the instructions are randomized, in comparison with when they are blocked. As such, they converge nicely with the conclusions reached by Petrusic (1992) localizing the SCE in the slowed accrual of evidence. In addition, Leth-Steensen and Marley's (2000) connectionist, instructional pathway interference model provides explicit mechanisms through which this slowing can be assumed to occur. That is, whenever the instructions are randomized, both instructions are available in memory and pathway interference is maximal, but whenever the instructions are blocked, activation of the irrelevant instructional pathway interference is considerably reduced, relative to the randomized condition, given that the irrelevant instruction is truly irrelevant.

In contrast, theories of the SCE that do not conceptualize decision processing in symbolic comparison in terms of the accumulation of evidence seemingly require the clearly ad hoc assumption that presenting the instructions in a randomized fashion invokes a mental load that limits the amount of cognitive resources that are available to the comparison process itself. Although the notion of limited cognitive resources typically has received general acceptance in the cognitive literature, it could also be argued that it is actually a rather ill-defined construct, whose use as a "catch-all" explanation for any number of load-type effects has greatly diminished its theoretical value.

In any case, we believe that the present results provide an additional empirical constraint that must now be part of the taxonomy of empirical constraints currently relevant to all theories of the symbolic comparison process. In addition, much more work, involving other kinds of instructional-based manipulations, is currently being undertaken by us (e.g., Shaki & Petrusic, 2003). Our hope is that the results of all of this work, taken together, will eventually provide a strong *set* of instructional-based empirical constraints that will allow for a better determination of the validity of each of the theories discussed herein.

REFERENCES

Audley, R. J., & Wallis, C. P. (1964). Response instructions and the speed of relative judgments: I. Some experiments on brightness discrimination. *British Journal of Psychology*, **55**, 59-73.

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 101-159). New York: Academic Press.

Banks, W. P., Clark, H. H., & Lucy, P. (1975). The locus of the semantic congruity effect in comparative judgments. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 35-47.

Banks, W. P., & Flora, J. (1977). Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 278-290.

Banks, W. P., Fujii, M. S., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 435-447.

Banks, W. P., & Root, M. (1979). Semantic congruity effects in judgments of loudness. *Perception & Psychophysics*, **26**, 133-142.

Birnbaum, M. H., & Jou, J. (1990). A theory of comparative response times and "difference" judgments. *Cognitive Psychology*, **22**, 184-210.

Cech, C. G. (1995). Is congruity due to encoding? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **5**, 1275-1288.

Cech, C. G., & Shoben, E. J. (1985). Context effects in symbolic magnitude comparisons. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 299-315.

Cech, C. G., Shoben, E. J., & Love, M. (1990). Multiple congruity effects in judgments of magnitude. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1142-1152.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, **97**, 332-361.

Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception & Psychophysics*, **45**, 557-566.

Holyoak, K. J. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology*, **10**, 203-243.

Jamieson, D. G., & Petrusic, W. M. (1975). Relational judgments with remembered stimuli. *Perception & Psychophysics*, **18**, 373-378.

Kosslyn, S. M., Murphy, G. L., Bemesderfer, M. E., & Feinstein, K. J. (1977). Category and continuum in mental comparisons. *Journal of Experimental Psychology: General*, **106**, 341-375.

Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, **107**, 62-100.

Link, S. W. (1990). Modeling imageless thought: The relative judgment theory of numerical comparisons. *Journal of Mathematical Psychology*, **34**, 2-41.

Link, S. W. (1992). *The wave theory of difference and similarity*. Hove, U.K.: Erlbaum.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, **22**, 1-35.

Los, S. A. (1996). On the origin of mixing costs: Exploring information processing in pure and mixed blocks of trials. *Acta Psychologica*, **94**, 145-188.

MacDonald, P. A., & Joordens, S. (2000). Investigating a memory-based account of negative priming: Support for selection-feature mismatch. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1478-1496.

Marks, D. F. (1972). Relative judgment: A phenomenon and a theory. *Perception & Psychophysics*, **11**, 156-160.

Marschark, M., & Paivio, A. (1979). Semantic congruity and lexical marking in symbolic comparisons: An expectancy hypothesis. *Memory & Cognition*, **7**, 175-184.

Marschark, M., & Paivio, A. (1981). Congruity and the perceptual comparison task. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 290-308.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, **90**, 227-234.

Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, **13**, 180-184.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, **106**, 226-254.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition*, **3**, 635-647.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, **116**, 220-244.

Petrusic, W. M. (1992). Semantic congruity effects and theories of the comparison process. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 962-986.

Petrusic, W. M., & Baranski, J. V. (1989). Semantic congruity effects in perceptual comparisons. *Perception & Psychophysics*, **45**, 439-452.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, **124**, 207-231.

Sanders, A. F. (1997). A summary of resource theories from a behavioral perspective. *Biological Psychology*, **45**, 5-18.

Schwarz, W., & Stein, F. (1998). On the temporal dynamics of digit comparison processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1275-1293.

Shaki, S., & Algom, D. (2002). The locus and nature of semantic congruity in symbolic comparison: Evidence from the Stroop effect. *Memory & Cognition*, **30**, 3-17.

Shaki, S., & Petrusic, W. M. (2003). Instruction interference and the semantic congruity effect. In B. Berglund & E. Borg (Eds.), *Fechner Day 2003: Proceedings of the Nineteenth Annual Meeting of the International Society for Psychophysics* (pp. 293-298). Larnaca Bay, Cyprus: International Society for Psychophysics.

Strayer, D. L., & Grison, S. (1999). Negative identity priming is contingent on stimulus repetition. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 24-38.

Wallis, C. P., & Audley, R. J. (1964). Response instructions and the speed of relative judgments: II. Pitch discrimination. *British Journal of Psychology*, **55**, 121-132.

Wood, T. J., & Milliken, B. (1998). Negative priming without ignoring. *Psychonomic Bulletin & Review*, **5**, 470-475.

NOTES

1. Responses were made on a second serial mouse with the roller ball disabled. SuperLab documentation states that "SuperLab accesses its buttons directly using the serial port and obtains 1 msec accuracy."

2. One aspect of presenting the comparative instructions in a randomly intermixed fashion, as compared with a blocked fashion, is that identical stimulus repetitions over consecutive trials (i.e., the same pair in the same left–right position with the same comparative instruction) would potentially be more likely in the blocked instruction case (Los, 1996). Furthermore, because responses to such identical repetitions could be made simply by quickly repeating the previous response, rather than by invoking any actual decision process, the mixing of such responses with actual comparison responses would then serve to attenuate both the SCE and the DE in the blocked instruction presentation conditions. However, note that this point is actually a nonissue in Experiment 1, because the stimulus set defined by the basic design was presented in a blocked fashion, which then precluded the possibility of exact stimulus repetitions.

APPENDIX A

Mean Response Times (RTs, in Milliseconds), Standard Deviations, and Numbers (*N*s) for Each Pair in the Design for Correct and Error Responses in the Blocked Condition

	Ant			Bee			Rat			Cat			Hog			Cow		
	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*
Ant				821.2	519.0	90	695.5	275.6	95	714.1	328.9	96	710.1	364.3	95	784.1	620.1	96
				777.7	429.1	6	552.0	–	1	–	–	0	2,441.0	–	1	–	–	0
				1,132.3	485.1	89	966.9	405.8	95	842.1	–	96	851.8	–	96	681.4	–	96
				1,128.1	508.7	7	527.0	–	1	–	–	0	–	–	0	–	–	0
Bee	832.9	396.7	83				1,104.3	1,089.9	93	799.1	266.2	94	900.3	620.6	96	763.9	375.0	96
	762.7	298.4	13				890.3	387.0	3	495.5	60.1	2	–	–	0	–	–	0
	1,176.6	416.5	89				1,071.8	408.9	95	1,043.4	1,099.2	95	854.8	312.5	96	688.2	264.1	95
	895.9	278.4	7				502.0	–	1	551.0	–	1	–	–	0	601.0	–	1
Rat	750.1	235.0	96	1,012.1	699.3	95				1,268.1	664.3	86	1,183.8	543.8	95	896.1	316.9	96
	–	–	0	424.0	–	1				756.7	210.5	10	879.0	–	1	–	–	0
	985.1	442.8	95	1,158.8	605.9	92				1,077.7	394.1	83	1,037.9	576.9	96	718.9	456.1	96
	459.0	–	1	1,365.3	1,249.4	4				869.4	251.3	13	–	–	0	–	–	0
Cat	781.4	507.3	95	939.9	627.9	94	1,235.7	568.9	86				1,286.1	531.0	91	925.4	384.4	96
	454.0	–	1	450.0	155.6	2	728.9	288.9	10				1,114.0	624.8	5	–	–	0
	918.1	339.1	95	987.3	359.8	95	1,122.5	494.8	91				1,102.8	700.3	80	692.5	237.7	95
	795.0	–	1	582.0	–	1	895.0	413.0	5				864.2	457.6	16	593.0	–	1
Hog	738.9	296.2	96	968.8	740.3	94	1,206.8	589.6	93	1,271.9	627.8	87				1,062.9	413.9	93
	–	–	0	646.5	154.9	2	503.0	21.9	3	695.8	291.2	9				1,070.0	665.2	3
	974.4	1,020.2	94	866.9	485.3	95	1,005.0	430.9	91	1,013.0	352.1	83				759.4	389.8	95
	655.0	62.2	2	798.0	–	1	725.4	341.9	5	1,024.4	881.6	13				364.0	–	1
Cow	640.8	195.9	96	759.2	340.6	95	932.1	481.1	96	1,035.2	489.3	96	1,025.4	463.8	91			
	–	–	0	541.0	–	1	–	–	0	–	–	0	1,079.4	683.3	5			
	713.9	408.4	96	666.8	313.4	96	669.2	254.8	96	675.2	251.1	96	707.9	347.8	95			
	–	–	0	–	–	0	–	–	0	–	–	0	548.0	–	1			

Note—The stimuli for the rows appeared on the left in each pair, and the stimuli over the columns appeared on the right. The entries in the first two rows in each cell arise with the instruction to choose the smaller animal in the pair, and the entries in the third and fourth rows arise with the instruction to choose the larger animal in the pair. Entries in the first and third rows in each cell correspond to correct responses, and entries in the second and fourth rows in each cell correspond to error responses.

APPENDIX B

Mean Response Times (RTs, in Milliseconds), Standard Deviations (*SD*s), and Numbers (*N*s) for Each Pair in the Design for Correct and Error Responses in the Randomized Condition

	Ant			Bee			Rat			Cat			Hog			Cow		
	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*	RT	*SD*	*N*
Ant				968.1	515.3	94	942.9	622.2	94	936.8	686.4	95	862.8	464.9	95	858.2	568.7	94
				2,007.0	613.8	2	907.5	457.4	2	1,071.0	–	1	348.0	–	1	597.0	144.3	2
				1,424.4	710.1	91	1,158.1	777.8	91	1,013.0	536.6	93	949.7	514.9	93	844.6	514.9	93
				1,088.8	298.7	5	997.4	453.9	5	811.0	108.7	3	994.7	315.2	3	989.3	516.8	3
Bee	1,062.5	630.7	90				1,096.4	500.7	90	1,087.7	827.4	96	1,010.4	591.2	95	1,013.8	838.8	95
	1,109.5	399.1	6				918.5	685.5	6	–	–	0	592.0	–	1	442.0	–	1
	1,428.2	686.3	87				1,313.4	657.2	90	1,212.8	678.6	94	1,026.6	641.4	93	809.7	431.1	94
	1,035.6	422.3	9				1,566.7	597.3	6	1,230.5	58.7	2	895.3	299.9	3	682.0	254.6	2
Rat	897.9	377.2	94	1,118.8	563.2	93				1,346.2	1,059.6	88	1,282.9	1,017.3	95	1,075.9	498.4	95
	639.5	14.8	2	1,920.0	968.1	3				1,250.1	239.5	8	1,413.0	–	1	1,794.0	–	1
	1,242.2	602.6	95	1,346.1	712.0	94				1,420.0	872.2	90	1,158.8	731.8	93	881.5	566.0	92
	651.0	–	1	617.0	69.3	2				963.3	362.9	6	1,110.7	835.4	3	948.8	146.5	4
Cat	954.6	563.0	92	1,032.9	488.0	95	1,373.9	1,133.6	86				1,508.6	771.7	88	1,148.8	667.9	94
	993.0	160.0	4	512.0	–	1	859.8	289.5	10				857.0	316.0	8	512.5	71.4	2
	1,141.9	660.6	93	1,178.1	751.6	94	1,303.4	552.9	93				1,188.5	468.0	86	890.8	717.2	94
	823.7	251.4	3	761.5	92.6	2	986.3	502.3	3				1,058.0	978.4	10	694.5	259.5	2
Hog	961.6	604.7	94	1,029.1	575.1	92	1,354.3	644.4	91	1,444.8	631.7	88				1,357.4	926.0	93
	2,493.0	2,225.9	2	761.8	148.6	4	1,193.6	498.4	5	968.4	474.9	8				1,085.7	487.6	3
	1,131.6	802.1	94	1,161.8	716.3	93	1,258.3	633.8	93	1,338.9	687.9	87				869.7	489.4	95
	754.5	89.8	2	484.0	96.2	3	822.0	234.2	3	1,532.0	1,171.9	9				1,653.0	–	1
Cow	808.9	366.7	95	921.3	523.3	92	1,014.6	511.7	95	1,116.2	502.1	93	1,366.2	1,133.6	89			
	920.0	–	1	1,228.5	358.5	4	814.0	–	1	986.0	607.7	3	1,938.0	1,631.7	7			
	816.8	470.0	94	819.4	467.4	94	810.1	482.3	95	864.9	436.2	92	867.8	466.5	95			
	803.0	540.2	2	1,062.0	16.9	2	543.0	–	1	647.3	280.6	4	805.0	–	1			

Note—The stimuli for the rows appeared on the left in each pair, and the stimuli over the columns appeared on the right. The entries in the first two rows in each cell arise with the instruction to choose the smaller animal in the pair, and the entries in the third and fourth rows arise with the instruction to choose the larger animal in the pair. Entries in the first and third rows in each cell correspond to correct responses, and entries in the second and fourth rows correspond to error responses.

(Manuscript received October 28, 2003;
revision accepted for publication January 16, 2005.)